GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2003, 01:11:33 ; Search time 178.416 Seconds
(without alignments)
10110.387 Million cell updates/sec

Title: US-09-434-382-28_COPY_21800_22600
Perfect score: 801
Sequence: 1 agtgcctgtcctggtatttt.....agcggaagctttgaccggat 801

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	800.6	100.0	26664	21	AAA60207	Human prostate can
	2	800.6	100.0	26664	24	AAS98942	Human prostate can
c	3	800.6	100.0	37959	22	AAK81763	Human immune/haema
	4	145.8	18.2	2481	24	AAS98916	Human prostate can
	5	145.8	18.2	2546	21	AAC76445	Human ORFX ORF2000
	6	145.8	18.2	2892	24	AAS99133	Gorilla ELAC2 cDNA
	7	145.8	18.2	2908	24	AAS99132	Chimpanzee ELAC2 c
	8	145.8	18.2	2958	21	AAA58453	Human prostate can
	9	145.8	18.2	2958	24	AAS98917	Human prostate can
	10	145.8	18.2	2976	22	AAH14250	Human cDNA sequenc
	11	145.8	18.2	2992	24	ABN59829	Novel human coding
	12	144.2	18.0	2478	21	AAA52810	Human sulphatase G
	13	141.6	17.7	1402	23	AAS72207	DNA encoding novel
	14	138.6	17.3	139	24	AAS98934	Human prostate can
	15	102.2	12.8	2470	24	AAS99131	Mouse ELAC2 cDNA.
	16	40.6	5.1	6370	24	ABL70568	Chemically treated
	17	40.6	5.1	6370	24	ABK31349	Signal transductio
	18	40.4	5.0	6121	24	ABL70151	Chemically treated
	19	40.4	5.0	6121	24	ABL34478	Human metastasis a
	20	40.4	5.0	6121	24	AAS61089	Human gene regulat
	21	40	5.0	10710	24	ABL32893	Human immune syste
c	22	39.4	4.9	840	24	ABL60745	Maize haemoglobin
	23	39.2	4.9	1152	23	ABK42630	Genomic sequence #
c	24	39.2	4.9	1473	21	AAZ54347	Neisseria meningit
	25	39.2	4.9	1655	22	ABA20093	Human nervous syst
	26	39.2	4.9	1655	23	ABK42631	Genomic sequence #
	27	38.8	4.8	1732	22	ABA06713	Human cDNA SEQ ID
	28	38.8	4.8	1732	22	AAS28782	Human immunoglobul
	29	38.8	4.8	3169	22	AAK94774	Human full-length
	30	38.8	4.8	6015	21	AAC75931	Human ORFX ORF1486
	31	38.8	4.8	15548	24	ABL34155	Human immune syste
	32	38	4.7	5981	24	ABL70484	Chemically treated
	33	38	4.7	5981	24	ABL34193	Human immune syste
	34	38	4.7	10286	22	AAS45308	Chemically pretrea
	35	38	4.7	10286	24	ABK28147	DNA transcription
c	36	37.6	4.7	1476	21	AAZ54345	Neisseria gonorrhe
	37	37.2	4.6	1899	22	AAF44678	Novel protein kina
	38	37.2	4.6	2232	23	ABL02765	Drosophila melanog
	39	37.2	4.6	4290	23	ABL02764	Drosophila melanog
	40	36.8	4.6	671	23	ABK41898	cDNA encoding nove
c	41	36.8	4.6	810	24	ABN92488	Staphylococcus epi
	42	36.8	4.6	1578	24	AAS62796	cDNA sequence #583
c	43	36.8	4.6	4248	22	AAH54305	S. epidermidis gen
c	44	36.6	4.6	19124	18	AAT72882	Plasmodium var-7 g
c	45	36.6	4.6	19124	21	AAZ98287	Plasmodium var-7 p

ALIGNMENTS

RESULT 1
AAA60207
ID AAA60207 standard; DNA; 26664 BP.
XX
AC AAA60207;
XX
DT 07-DEC-2000 (first entry)
XX
DE Human prostate cancer predisposing gene HPC2 genomic sequence.
XX
KW Human; prostate cancer predisposing gene; HPC2; chromosome 17p;
KW gene therapy; peptide therapy; drug design; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 910..26039
FT /*tag= a
FT /product= "HPC2"
FT /note= "this sequence contains introns"
FT /transl_except= (pos:23892..23895,aa:Glu)
FT exon 910..1154
FT /*tag= b
FT /number= 1
FT exon 1736..1786
FT /*tag= c
FT /number= 2
FT exon 1925..1995
FT /*tag= d
FT /number= 3
FT exon 3025..3089
FT /*tag= e

```
FT                   /number= 4
FT   exon            4361..4418
FT                   /*tag=  f
FT                   /number= 5
FT   exon            5582..5650
FT                   /*tag=  g
FT                   /number= 6
FT   exon            7075..7194
FT                   /*tag=  h
FT                   /number= 7
FT   exon            8186..8244
FT                   /*tag=  i
FT                   /number= 8
FT   exon            12878..12936
FT                   /*tag=  j
FT                   /number= 9
FT   exon            13032..13104
FT                   /*tag=  k
FT                   /number= 10
FT   exon            13756..13868
FT                   /*tag=  l
FT                   /number= 11
FT   exon            15283..15378
FT                   /*tag=  m
FT                   /number= 12
FT   exon            16278..16416
FT                   /*tag=  n
FT                   /number= 13
FT   exon            16498..16583
FT                   /*tag=  o
FT                   /number= 14
FT   exon            18583..18701
FT                   /*tag=  p
FT                   /number= 15
FT   exon            20349..20445
FT                   /*tag=  q
FT                   /number= 16
FT   exon            22172..22310
FT                   /*tag=  r
FT                   /number= 17
FT   exon            22879..22917
FT                   /*tag=  s
FT                   /number= 18
FT   exon            23045..23154
FT                   /*tag=  t
FT                   /number= 19
FT   exon            23795..23895
FT                   /*tag=  u
FT                   /number= 20
FT   exon            23973..24093
FT                   /*tag=  v
FT                   /number= 21
FT   exon            24354..24432
FT                   /*tag=  w
FT                   /number= 22
FT   exon            25026..25170
FT                   /*tag=  x
FT                   /number= 23
FT   exon            25812..26036
FT                   /*tag=  y
FT                   /number= 24
FT   polyA_signal    26447..26452
FT                   /*tag=  z
XX
PN   WO200027864-A1.
XX
PD   18-MAY-2000.
XX
PF   05-NOV-1999;   99WO-US26055.
XX
PR   06-NOV-1998;   98US-0107468.
XX
PA   (MYRI-) MYRIAD GENETICS INC.
XX
PI   Tavtigian SV,  Teng DHF,  Simard J,  Rommens JM;
XX
DR   WPI; 2000-376481/32.
DR   P-PSDB; AAB07228.
XX
PT   Human prostate cancer (HPC)2 nucleic acids, polypeptides, and
PT   antibodies, useful for treatment and diagnosis of prostate cancer -
XX
PS   Claim 3; Page 108-122; 157pp; English.
XX
CC   The present sequence is the genomic sequence of the human prostate
CC   cancer predisposing gene HPC2, which is found on chromosome 17p. Some
CC   alleles of this gene cause a predisposition to cancer, particularly
CC   prostate cancer. This gene and its protein can be used in peptide and
CC   gene therapy for cancer patients, as well as being useful as diagnostic
CC   tools (both for cancer sufferers and those with a predisposition to the
CC   disease) and in the production of cancer drugs. This sequence was
CC   isolated by cloning and sequencing the region of the genome which
CC   appeared to cause a predisposition to prostate cancer.
XX
SQ   Sequence 26664 BP; 6173 A; 6300 C; 6519 G; 7661 T; 11 other;

  Query Match           100.0%;  Score 800.6;  DB 21;  Length 26664;
  Best Local Similarity  100.0%;  Pred. No. 6.4e-222;
  Matches  801;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AGTGCCTGTCCTGGTATTTTCAACAGAGGCTGTGGCCACAGTCAATCTGCATGGTCAGAT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   21800 AGTGCCTGTCCTGGTATTTTCAACAGAGGCTGTGGCCACAGTCAATCTGCATGGTCAGAT 21859

Qy      61 TCATTGTTAGGACTAAATGCTTTAAGCCTCCTATAAACTTTTTTTTTTTTTTTTTGATGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   21860 TCATTGTTAGGACTAAATGCTTTAAGCCTCCTATAAACTTTTTTTTTTTTTTTTTGATGC 21919

Qy     121 CCAGCCTTTGTGTAAGTCTACTTGAAAGGGTTTCAGGGTTCCATGGATACTTCTTTGCTA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   21920 CCAGCCTTTGTGTAAGTCTACTTGAAAGGGTTTCAGGGTTCCATGGATACTTCTTTGCTA 21979

Qy     181 TAAAGAGGATGACACATGTAAAATCACCTTTATGGTTAAATTAATTGGCTTTTATATTAG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   21980 TAAAGAGGATGACACATGTAAAATCACCTTTATGGTTAAATTAATTGGCTTTTATATTAG 22039

Qy     241 CTCCTCAAAGCAAAGCAGGAGAGACAGAAATTTCTGCAGTTGCTTCTTGGTCCTGTCCAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   22040 CTCCTCAAAGCAAAGCAGGAGAGACAGAAATTTCTGCAGTTGCTTCTTGGTCCTGTCCAA 22099

Qy     301 AGCAGACATCAGCCTCTGAACCATCAGCAGTCTTCCTAGTGGCAGTGACTCTCTTCCTCT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   22100 AGCAGACATCAGCCTCTGAACCATCAGCAGTCTTCCTAGTGGCAGTGACTCTCTTCCTCT 22159

Qy     361 TCTCTTCTGCAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACRTTTGGGCA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   22160 TCTCTTCTGCAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACRTTTGGGCA 22219

Qy     421 GCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCTGCTGTGTT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   22220 GCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCTGCTGTGTT 22279

Qy     481 TGTGTCCCACCTGCACGCAGATCACCACACGGTGAGTGTTGGGCTGGACCACAAAGCTGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   22280 TGTGTCCCACCTGCACGCAGATCACCACACGGTGAGTGTTGGGCTGGACCACAAAGCTGG 22339

Qy     541 AGCCTGGAGGAGGCACTGCCACGTTGAGTTGGCCCTTTGGCTGCGTCTTTTCCTCCGCTT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   22340 AGCCTGGAGGAGGCACTGCCACGTTGAGTTGGCCCTTTGGCTGCGTCTTTTCCTCCGCTT 22399

Qy     601 CCAAACTTGCCCAGAGCTTTTGTTACTCATCTCTGGCTAGGAAATGGTTTTTTGCAAAAC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   22400 CCAAACTTGCCCAGAGCTTTTGTTACTCATCTCTGGCTAGGAAATGGTTTTTTGCAAAAC 22459

Qy     661 TCAACATAGTCCTTCTGCGCCACAAGAATGTCTTCTCTTCCTGTTCAGTTCCTTTCCTGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  22460 TCAACATAGTCCTTCTGCGCCACAAGAATGTCTTCTCTTCCTGTTCAGTTCCTTTCCTGC 22519

Qy    721 AGCAGGACAGGTTTGAGTTTACCCAGCCTTCCTTGAGTCTTGAATCTCACACGGCCTGCT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  22520 AGCAGGACAGGTTTGAGTTTACCCAGCCTTCCTTGAGTCTTGAATCTCACACGGCCTGCT 22579

Qy    781 CAGCGGAAGCTTTGACCGGAT 801
          |||||||||||||||||||||
Db  22580 CAGCGGAAGCTTTGACCGGAT 22600
```

RESULT 2
AAS98942
ID AAS98942 standard; DNA; 26664 BP.
XX
AC AAS98942;
XX
DT 12-MAR-2002 (first entry)
XX
DE Human prostate cancer predisposing gene (HPC2) genomic DNA.
XX
KW Human; mouse; HPC2; prostate cancer; neoplastic growth; cytostatic; ds;
KW gene therapy; prostate cancer predisposing gene.
XX
OS Homo sapiens.
XX
PN WO200185911-A2.
XX
PD 15-NOV-2001.
XX
PF 07-MAY-2001; 2001WO-US14602.
XX
PR 05-MAY-2000; 2000US-0564805.
XX
PA (MYRI-) MYRIAD GENETICS INC.
PA (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI Tavtigian SV, Teng DHF, Simard J, Rommens JM;
XX
DR WPI; 2002-066599/09.
XX
PT Novel nucleic acid sequence encoding HPC2 polypeptide, which is marker
PT for prostate cancer, is useful in gene therapy techniques to restore
PT HPC2 normal levels by which neoplastic growth is suppressed in
PT recipient cell -
XX
PS Claim 3; Page 143-157; 239pp; English.
XX
CC The invention relates to a human prostate cancer predisposing gene coding
CC for an HPC2 polypeptide. The DNA and protein sequences are useful as
CC diagnostic reagents for identifying a mutant HPC2 nucleotide sequence in
CC a suspected mutant HPC2 allele by comparing the sequence of the suspected
CC mutant HPC2 allele with a wild-type HPC2 sequence. The sequences are also
CC useful for detecting an alteration in HPC2, where the alteration is
CC associated with cancer in a human. The method involves analysing an HPC2
CC gene or an HPC2 gene expression product from a tissue of the human. The
CC HPC2 gene is useful as a marker for prostate cancer and can be used in
CC gene therapy techniques to suppress neoplastic growth of recipient cells
CC which carry the mutant HPC2 allele. The sequences represent DNA encoding
CC human and mouse HPC2 and fragments of HPC2.
XX
SQ Sequence 26664 BP; 6173 A; 6300 C; 6519 G; 7661 T; 11 other;

```
  Query Match            100.0%;  Score 800.6;  DB 24;  Length 26664;
  Best Local Similarity  100.0%;  Pred. No. 6.4e-222;
  Matches  801;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AGTGCCTGTCCTGGTATTTTCAACAGAGGCTGTGGCCACAGTCAATCTGCATGGTCAGAT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21800 AGTGCCTGTCCTGGTATTTTCAACAGAGGCTGTGGCCACAGTCAATCTGCATGGTCAGAT 21859

Qy     61 TCATTGTTAGGACTAAATGCTTTAAGCCTCCTATAAACTTTTTTTTTTTTTTTTTGATGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21860 TCATTGTTAGGACTAAATGCTTTAAGCCTCCTATAAACTTTTTTTTTTTTTTTTTGATGC 21919

Qy    121 CCAGCCTTTGTGTAAGTCTACTTGAAAGGGTTTCAGGGTTCCATGGATACTTCTTTGCTA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21920 CCAGCCTTTGTGTAAGTCTACTTGAAAGGGTTTCAGGGTTCCATGGATACTTCTTTGCTA 21979

Qy    181 TAAAGAGGATGACACATGTAAAATCACCTTTATGGTTAAATTAATTGGCTTTTATATTAG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21980 TAAAGAGGATGACACATGTAAAATCACCTTTATGGTTAAATTAATTGGCTTTTATATTAG 22039

Qy    241 CTCCTCAAAGCAAAGCAGGAGAGACAGAAATTTCTGCAGTTGCTTCTTGGTCCTGTCCAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  22040 CTCCTCAAAGCAAAGCAGGAGAGACAGAAATTTCTGCAGTTGCTTCTTGGTCCTGTCCAA 22099

Qy    301 AGCAGACATCAGCCTCTGAACCATCAGCAGTCTTCCTAGTGGCAGTGACTCTCTTCCTCT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  22100 AGCAGACATCAGCCTCTGAACCATCAGCAGTCTTCCTAGTGGCAGTGACTCTCTTCCTCT 22159

Qy    361 TCTCTTCTGCAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACRTTTGGGCA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  22160 TCTCTTCTGCAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACRTTTGGGCA 22219

Qy    421 GCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCTGCTGTGTT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  22220 GCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCTGCTGTGTT 22279

Qy    481 TGTGTCCCACCTGCACGCAGATCACCACACGGTGAGTGTTGGGCTGGACCACAAAGCTGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  22280 TGTGTCCCACCTGCACGCAGATCACCACACGGTGAGTGTTGGGCTGGACCACAAAGCTGG 22339

Qy    541 AGCCTGGAGGAGGCACTGCCACGTTGAGTTGGCCCTTTGGCTGCGTCTTTTCCTCCGCTT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  22340 AGCCTGGAGGAGGCACTGCCACGTTGAGTTGGCCCTTTGGCTGCGTCTTTTCCTCCGCTT 22399

Qy    601 CCAAACTTGCCCAGAGCTTTTGTTACTCATCTCTGGCTAGGAAATGGTTTTTTGCAAAAC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  22400 CCAAACTTGCCCAGAGCTTTTGTTACTCATCTCTGGCTAGGAAATGGTTTTTTGCAAAAC 22459

Qy    661 TCAACATAGTCCTTCTGCGCCACAAGAATGTCTTCTCTTCCTGTTCAGTTCCTTTCCTGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  22460 TCAACATAGTCCTTCTGCGCCACAAGAATGTCTTCTCTTCCTGTTCAGTTCCTTTCCTGC 22519

Qy    721 AGCAGGACAGGTTTGAGTTTACCCAGCCTTCCTTGAGTCTTGAATCTCACACGGCCTGCT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  22520 AGCAGGACAGGTTTGAGTTTACCCAGCCTTCCTTGAGTCTTGAATCTCACACGGCCTGCT 22579

Qy    781 CAGCGGAAGCTTTGACCGGAT 801
          |||||||||||||||||||||
Db  22580 CAGCGGAAGCTTTGACCGGAT 22600
```

RESULT 3
AAK81763/c
ID AAK81763 standard; DNA; 37959 BP.
XX
AC AAK81763;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:36575.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX

```
PS   Disclosure; SEQ ID NO 36575; 3071pp + Sequence Listing; English.
XX
CC   AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC   amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC   activity, and can be used in gene therapy and vaccine production. (I)
CC   proteins and polynucleotides may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate (I) expression. For
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
CC   supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting
CC   the nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention.
XX
SQ   Sequence 37959 BP; 10440 A; 9397 C; 9111 G; 9011 T; 0 other;

  Query Match             100.0%;  Score 800.6;  DB 22;  Length 37959;
  Best Local Similarity   99.9%;  Pred. No. 7.7e-222;
  Matches  800;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AGTGCCTGTCCTGGTATTTTCAACAGAGGCTGTGGCCACAGTCAATCTGCATGGTCAGAT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   16995 AGTGCCTGTCCTGGTATTTTCAACAGAGGCTGTGGCCACAGTCAATCTGCATGGTCAGAT 16936

Qy      61 TCATTGTTAGGACTAAATGCTTTAAGCCTCCTATAAACTTTTTTTTTTTTTTTTTGATGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   16935 TCATTGTTAGGACTAAATGCTTTAAGCCTCCTATAAACTTTTTTTTTTTTTTTTTGATGC 16876

Qy     121 CCAGCCTTTGTGTAAGTCTACTTGAAAGGGTTTCAGGGTTCCATGGATACTTCTTTGCTA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   16875 CCAGCCTTTGTGTAAGTCTACTTGAAAGGGTTTCAGGGTTCCATGGATACTTCTTTGCTA 16816

Qy     181 TAAAGAGGATGACACATGTAAAATCACCTTTATGGTTAAATTAATTGGCTTTTATATTAG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   16815 TAAAGAGGATGACACATGTAAAATCACCTTTATGGTTAAATTAATTGGCTTTTATATTAG 16756

Qy     241 CTCCTCAAAGCAAAGCAGGAGAGACAGAAATTTCTGCAGTTGCTTCTTGGTCCTGTCCAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   16755 CTCCTCAAAGCAAAGCAGGAGAGACAGAAATTTCTGCAGTTGCTTCTTGGTCCTGTCCAA 16696

Qy     301 AGCAGACATCAGCCTCTGAACCATCAGCAGTCTTCCTAGTGGCAGTGACTCTCTTCCTCT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   16695 AGCAGACATCAGCCTCTGAACCATCAGCAGTCTTCCTAGTGGCAGTGACTCTCTTCCTCT 16636

Qy     361 TCTCTTCTGCAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACRTTTGGGCA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db   16635 TCTCTTCTGCAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACATTTGGGCA 16576

Qy     421 GCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCTGCTGTGTT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   16575 GCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCTGCTGTGTT 16516

Qy     481 TGTGTCCCACCTGCACGCAGATCACCACACGGTGAGTGTTGGGCTGGACCACAAAGCTGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   16515 TGTGTCCCACCTGCACGCAGATCACCACACGGTGAGTGTTGGGCTGGACCACAAAGCTGG 16456

Qy     541 AGCCTGGAGGAGGCACTGCCACGTTGAGTTGGCCCTTTGGCTGCGTCTTTTCCTCCGCTT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   16455 AGCCTGGAGGAGGCACTGCCACGTTGAGTTGGCCCTTTGGCTGCGTCTTTTCCTCCGCTT 16396

Qy     601 CCAAACTTGCCCAGAGCTTTTGTTACTCATCTCTGGCTAGGAAATGGTTTTTTGCAAAAC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   16395 CCAAACTTGCCCAGAGCTTTTGTTACTCATCTCTGGCTAGGAAATGGTTTTTTGCAAAAC 16336

Qy     661 TCAACATAGTCCTTCTGCGCCACAAGAATGTCTTCTCTTCCTGTTCAGTTCCTTTCCTGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   16335 TCAACATAGTCCTTCTGCGCCACAAGAATGTCTTCTCTTCCTGTTCAGTTCCTTTCCTGC 16276

Qy     721 AGCAGGACAGGTTTGAGTTTACCCAGCCTTCCTTGAGTCTTGAATCTCACACGGCCTGCT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   16275 AGCAGGACAGGTTTGAGTTTACCCAGCCTTCCTTGAGTCTTGAATCTCACACGGCCTGCT 16216

Qy     781 CAGCGGAAGCTTTGACCGGAT 801
          |||||||||||||||||||||
Db   16215 CAGCGGAAGCTTTGACCGGAT 16195


RESULT 4
AAS98916
ID   AAS98916 standard; cDNA; 2481 BP.
XX
AC   AAS98916;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Human prostate cancer predisposing gene (HPC2) cDNA coding sequence.
XX
KW   Human; mouse; HPC2; prostate cancer; neoplastic growth; cytostatic; ss;
KW   gene therapy; prostate cancer predisposing gene; chimpanzee; gorilla;
KW   sequencing primer; PCR primer.
XX
OS   Homo sapiens.
XX
PN   WO200185911-A2.
XX
PD   15-NOV-2001.
XX
PF   07-MAY-2001; 2001WO-US14602.
XX
PR   05-MAY-2000; 2000US-0564805.
XX
PA   (MYRI-) MYRIAD GENETICS INC.
PA   (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI   Tavtigian SV,  Teng DHF,  Simard J,  Rommens JM;
XX
DR   WPI; 2002-066599/09.
DR   P-PSDB; AAU73586.
XX
PT   Novel nucleic acid sequence encoding HPC2 polypeptide, which is marker
PT   for prostate cancer, is useful in gene therapy techniques to restore
PT   HPC2 normal levels by which neoplastic growth is suppressed in
PT   recipient cell  -
XX
PS   Claim 2; Page 128-131; 239pp; English.
XX
CC   The invention relates to a human prostate cancer predisposing gene coding
CC   for an HPC2 polypeptide. The DNA and protein sequences are useful as
CC   diagnostic reagents for identifying a mutant HPC2 nucleotide sequence in
CC   a suspected mutant HPC2 allele by comparing the sequence of the suspected
CC   mutant HPC2 allele with a wild-type HPC2 sequence. The sequences are also
CC   useful for detecting an alteration in HPC2, where the alteration is
CC   associated with cancer in a human. The method involves analysing an HPC2
CC   gene or an HPC2 gene expression product from a tissue of the human. The
CC   HPC2 gene is useful as a marker for prostate cancer and can be used in
CC   gene therapy techniques to suppress neoplastic growth of recipient cells
CC   which carry the mutant HPC2 allele. The sequences represent primers used
CC   in the methods of the invention, cDNA encoding human and mouse HPC2 and
CC   cDNA encoding HPC2 paralogues and orthologues.
XX
SQ   Sequence 2481 BP; 588 A; 686 C; 711 G; 496 T; 0 other;

  Query Match             18.2%;  Score 145.8;  DB 24;  Length 2481;
  Best Local Similarity   89.1%;  Pred. No. 8.4e-32;
  Matches  156;  Conservative    1;  Mismatches   18;  Indels    0;  Gaps    0;

Qy     338 AGTGGCAGTGACTCTCTTCCTCTTCTCTTCTGCAGCCCCGACACGTCTCTGCTACTGGAC 397
           | | | | || |  |    | ||| ||  |  ||||||||||||||||||||||||
Db    1486 ATTCGAAATGTCAGTGCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGAC 1545
```

```
Qy     398 TGTGGTGAGGGCACRTTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTC 457
           ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db    1546 TGTGGTGAGGGCACATTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTC 1605

Qy     458 CTGGGCACCCTGGCTGCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGG 512
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1606 CTGGGCACCCTGGCTGCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGG 1660


RESULT 5
AAC76445
ID   AAC76445 standard; cDNA; 2546 BP.
XX
AC   AAC76445;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human ORFX ORF2000 polynucleotide sequence SEQ ID NO:3999.
XX
KW   Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW   vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW   anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW   immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW   hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW   antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW   antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW   neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW   cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW   cholesterol ester storage; systemic lupus erythematosus; infection;
KW   severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW   allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW   bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW   thrombosis; contraceptive; ss.
XX
OS   Homo sapiens.
XX
PN   WO200058473-A2.
XX
PD   05-OCT-2000.
XX
PF   31-MAR-2000; 2000WO-US08621.
XX
PR   31-MAR-1999;   99US-0127607.
PR   02-APR-1999;   99US-0127636.
PR   05-APR-1999;   99US-0127728.
PR   30-MAR-2000; 2000US-0540763.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2000-602362/57.
DR   P-PSDB; AAB42236.
XX
PT   Novel nucleic acids and peptides derived from open reading frame X,
PT   useful for treating e.g. cancers, proliferative disorders,
PT   neurodegenerative disorders and cardiovascular disease -
XX
PS   Claim 5; Page 3179-3180; 5507pp; English.
XX
CC   AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC   which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC   sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC   antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC   osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC   immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC   antidiabetic; hypotensive; dermatological; immunosuppressive;
CC   antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC   antithyroid; and antianaemic. The sequences can be used for determining
CC   the presence of or predisposition to, or preventing or treating
CC   pathological conditions associated with an ORFX-associated disorder. The
CC   nucleic acids can be used to express ORFX proteins in gene therapy
CC   vectors. The proteins and nucleic acids may be used to treat cancers,
CC   proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC   graft vs host disease, cardiovascular disease, diabetes mellitus,
CC   hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC   erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC   bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC   allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC   nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC   coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ   Sequence 2546 BP; 652 A; 643 C; 686 G; 564 T; 1 other;

  Query Match             18.2%;  Score 145.8;  DB 21;  Length 2546;
  Best Local Similarity   89.1%;  Pred. No. 8.6e-32;
  Matches  156;  Conservative    1;  Mismatches   18;  Indels    0;  Gaps    0;

Qy     338 AGTGGCAGTGACTCTCTTCCTCTTCTCTTCTGCAGCCCCGACACGTCTCTGCTACTGGAC 397
           |·| | | || |  |    | ||| ||  |  ||||||||||||||||||||||||||
Db    1084 ATTCGAAATGTCAGTGCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGAC 1143

Qy     398 TGTGGTGAGGGCACRTTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTC 457
           |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db    1144 TGTGGTGAGGGCACATTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTC 1203

Qy     458 CTGGGCACCCTGGCTGCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGG 512
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1204 CTGGGCACCCTGGCTGCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGG 1258


RESULT 6
AAS99133
ID   AAS99133 standard; cDNA; 2892 BP.
XX
AC   AAS99133;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Gorilla ELAC2 cDNA.
XX
KW   Human; mouse; HPC2; prostate cancer; neoplastic growth; cytostatic; ss;
KW   gene therapy; prostate cancer predisposing gene; chimpanzee; gorilla;
KW   sequencing primer; PCR primer.
XX
OS   Gorilla gorilla.
XX
PN   WO200185911-A2.
XX
PD   15-NOV-2001.
XX
PF   07-MAY-2001; 2001WO-US14602.
XX
PR   05-MAY-2000; 2000US-0564805.
XX
PA   (MYRI-) MYRIAD GENETICS INC.
PA   (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI   Tavtigian SV,  Teng DHF,  Simard J,  Rommens JM;
XX
DR   WPI; 2002-066599/09.
DR   P-PSDB; AAU73593.
XX
PT   Novel nucleic acid sequence encoding HPC2 polypeptide, which is marker
PT   for prostate cancer, is useful in gene therapy techniques to restore
PT   HPC2 normal levels by which neoplastic growth is suppressed in
PT   recipient cell  -
XX
PS   Claim 92; Page 204-207; 239pp; English.
XX
CC   The invention relates to a human prostate cancer predisposing gene coding
CC   for an HPC2 polypeptide. The DNA and protein sequences are useful as
CC   diagnostic reagents for identifying a mutant HPC2 nucleotide sequence in
CC   a suspected mutant HPC2 allele by comparing the sequence of the suspected
CC   mutant HPC2 allele with a wild-type HPC2 sequence. The sequences are also
CC   useful for detecting an alteration in HPC2, where the alteration is
```

```
CC   associated with cancer in a human. The method involves analysing an HPC2
CC   gene or an HPC2 gene expression product from a tissue of the human. The
CC   HPC2 gene is useful as a marker for prostate cancer and can be used in
CC   gene therapy techniques to suppress neoplastic growth of recipient cells
CC   which carry the mutant HPC2 allele. The sequences represent primers used
CC   in the methods of the invention, cDNA encoding human and mouse HPC2 and
CC   cDNA encoding HPC2 paralogues and orthologues.
XX
SQ   Sequence 2892 BP; 704 A; 787 C; 815 G; 586 T; 0 other;

  Query Match             18.2%;  Score 145.8;  DB 24;  Length 2892;
  Best Local Similarity   89.1%;  Pred. No. 9.1e-32;
  Matches  156;  Conservative    1;  Mismatches   18;  Indels    0;  Gaps    0;

Qy     338 AGTGGCAGTGACTCTCTTCCTCTTCTCTTCTGCAGCCCCGACACGTCTCTGCTACTGGAC 397
           | | | | || |  |    | ||| ||  |   ||||||||||||||||||||||||||
Db    1486 ATTCGAAATGTCAGTGCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGAC 1545

Qy     398 TGTGGTGAGGGCACRTTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTC 457
           |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db    1546 TGTGGTGAGGGCACGTTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTC 1605

Qy     458 CTGGGCACCCTGGCTGCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGG 512
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1606 CTGGGCACCCTGGCTGCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGG 1660


RESULT 7
AAS99132
ID   AAS99132 standard; cDNA; 2908 BP.
XX
AC   AAS99132;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Chimpanzee ELAC2 cDNA.
XX
KW   Human; mouse; HPC2; prostate cancer; neoplastic growth; cytostatic; ss;
KW   gene therapy; prostate cancer predisposing gene; chimpanzee; gorilla;
KW   sequencing primer; PCR primer.
XX
OS   Pan troglodytes.
XX
PN   WO200185911-A2.
XX
PD   15-NOV-2001.
XX
PF   07-MAY-2001; 2001WO-US14602.
XX
PR   05-MAY-2000; 2000US-0564805.
XX
PA   (MYRI-) MYRIAD GENETICS INC.
PA   (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI   Tavtigian SV,  Teng DHF,  Simard J,  Rommens JM;
XX
DR   WPI; 2002-066599/09.
DR   P-PSDB; AAU73592.
XX
PT   Novel nucleic acid sequence encoding HPC2 polypeptide, which is marker
PT   for prostate cancer, is useful in gene therapy techniques to restore
PT   HPC2 normal levels by which neoplastic growth is suppressed in
PT   recipient cell  -
XX
PS   Claim 87; Page 198-201; 239pp; English.
XX
CC   The invention relates to a human prostate cancer predisposing gene coding
CC   for an HPC2 polypeptide. The DNA and protein sequences are useful as
CC   diagnostic reagents for identifying a mutant HPC2 nucleotide sequence in
CC   a suspected mutant HPC2 allele by comparing the sequence of the suspected
CC   mutant HPC2 allele with a wild-type HPC2 sequence. The sequences are also
CC   useful for detecting an alteration in HPC2, where the alteration is
CC   associated with cancer in a human. The method involves analysing an HPC2
CC   gene or an HPC2 gene expression product from a tissue of the human. The
CC   HPC2 gene is useful as a marker for prostate cancer and can be used in
CC   gene therapy techniques to suppress neoplastic growth of recipient cells
CC   which carry the mutant HPC2 allele. The sequences represent primers used
CC   in the methods of the invention, cDNA encoding human and mouse HPC2 and
CC   cDNA encoding HPC2 paralogues and orthologues.
XX
SQ   Sequence 2908 BP; 712 A; 788 C; 819 G; 589 T; 0 other;

  Query Match             18.2%;  Score 145.8;  DB 24;  Length 2908;
  Best Local Similarity   89.1%;  Pred. No. 9.1e-32;
  Matches  156;  Conservative    1;  Mismatches   18;  Indels    0;  Gaps    0;

Qy     338 AGTGGCAGTGACTCTCTTCCTCTTCTCTTCTGCAGCCCCGACACGTCTCTGCTACTGGAC 397
           | | | | || |  |    | ||| ||  |   ||||||||||||||||||||||||||
Db    1486 ATTCGAAATGTCAGTGCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGAC 1545

Qy     398 TGTGGTGAGGGCACRTTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTC 457
           |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db    1546 TGTGGTGAGGGCACGTTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTC 1605

Qy     458 CTGGGCACCCTGGCTGCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGG 512
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1606 CTGGGCACCCTGGCTGCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGG 1660


RESULT 8
AAA58453
ID   AAA58453 standard; cDNA; 2958 BP.
XX
AC   AAA58453;
XX
DT   07-DEC-2000  (first entry)
XX
DE   Human prostate cancer predisposing gene HPC2 coding sequence.
XX
KW   Human; prostate cancer predisposing gene; HPC2; chromosome 17p;
KW   gene therapy; peptide therapy; drug design; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             51..2531
FT                   /*tag=  a
FT                   /product= "HPC2"
XX
PN   WO200027864-A1.
XX
PD   18-MAY-2000.
XX
PF   05-NOV-1999;   99WO-US26055.
XX
PR   06-NOV-1998;   98US-0107468.
XX
PA   (MYRI-) MYRIAD GENETICS INC.
XX
PI   Tavtigian SV,  Teng DHF,  Simard J,  Rommens JM;
XX
DR   WPI; 2000-376481/32.
DR   P-PSDB; AAB07228.
XX
PT   Human prostate cancer (HPC)2 nucleic acids, polypeptides, and
PT   antibodies, useful for treatment and diagnosis of prostate cancer  -
XX
PS   Claim 3; Page 98-100; 157pp; English.
XX
CC   The present sequence is the coding sequence of the human prostate
CC   cancer predisposing gene HPC2, which is found on chromosome 17p. Some
CC   alleles of this gene cause a predisposition to cancer, particularly
CC   prostate cancer. This gene and its protein can be used in peptide and
CC   gene therapy for cancer patients, as well as being useful as diagnostic
CC   tools (both for cancer sufferers and those with a predisposition to the
CC   disease) and in the production of cancer drugs. This sequence was
```

```
CC   isolated by cloning and sequencing the region of the genome which
CC   appeared to cause a predisposition to prostate cancer.
XX
SQ   Sequence 2958 BP; 707 A; 805 C; 848 G; 598 T; 0 other;

  Query Match             18.2%;  Score 145.8;  DB 21;  Length 2958;
  Best Local Similarity   89.1%;  Pred. No. 9.2e-32;
  Matches  156;  Conservative    1;  Mismatches   18;  Indels    0;  Gaps    0;

Qy     338 AGTGGCAGTGACTCTCTTCCTCTTCTCTTCTGCAGCCCCGACACGTCTCTGCTACTGGAC 397
           | | | | || |  |    | ||| ||  |   ||||||||||||||||||||||||
Db    1536 ATTCGAAATGTCAGTGCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGAC 1595

Qy     398 TGTGGTGAGGGCACRTTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTC 457
           |||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db    1596 TGTGGTGAGGGCACATTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTC 1655

Qy     458 CTGGGCACCCTGGCTGCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGG 512
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1656 CTGGGCACCCTGGCTGCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGG 1710


RESULT 9
AAS98917
ID   AAS98917 standard; cDNA; 2958 BP.
XX
AC   AAS98917;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Human prostate cancer predisposing gene (HPC2) extended cDNA.
XX
KW   Human; mouse; HPC2; prostate cancer; neoplastic growth; cytostatic; ss;
KW   gene therapy; prostate cancer predisposing gene; chimpanzee; gorilla;
KW   sequencing primer; PCR primer.
XX
OS   Homo sapiens.
XX
PN   WO200185911-A2.
XX
PD   15-NOV-2001.
XX
PF   07-MAY-2001; 2001WO-US14602.
XX
PR   05-MAY-2000; 2000US-0564805.
XX
PA   (MYRI-) MYRIAD GENETICS INC.
PA   (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI   Tavtigian SV,  Teng DHF,  Simard J,  Rommens JM;
XX
DR   WPI; 2002-066599/09.
DR   P-PSDB; AAU73586.
XX
PT   Novel nucleic acid sequence encoding HPC2 polypeptide, which is marker
PT   for prostate cancer, is useful in gene therapy techniques to restore
PT   HPC2 normal levels by which neoplastic growth is suppressed in
PT   recipient cell  -
XX
PS   Claim 3; Page 134-136; 239pp; English.
XX
CC   The invention relates to a human prostate cancer predisposing gene coding
CC   for an HPC2 polypeptide. The DNA and protein sequences are useful as
CC   diagnostic reagents for identifying a mutant HPC2 nucleotide sequence in
CC   a suspected mutant HPC2 allele by comparing the sequence of the suspected
CC   mutant HPC2 allele with a wild-type HPC2 sequence. The sequences are also
CC   useful for detecting an alteration in HPC2, where the alteration is
CC   associated with cancer in a human. The method involves analysing an HPC2
CC   gene or an HPC2 gene expression product from a tissue of the human. The
CC   HPC2 gene is useful as a marker for prostate cancer and can be used in
CC   gene therapy techniques to suppress neoplastic growth of recipient cells
CC   which carry the mutant HPC2 allele. The sequences represent primers used
CC   in the methods of the invention, cDNA encoding human and mouse HPC2 and
CC   cDNA encoding HPC2 paralogues and orthologues.
XX
SQ   Sequence 2958 BP; 707 A; 805 C; 848 G; 598 T; 0 other;

  Query Match             18.2%;  Score 145.8;  DB 24;  Length 2958;
  Best Local Similarity   89.1%;  Pred. No. 9.2e-32;
  Matches  156;  Conservative    1;  Mismatches   18;  Indels    0;  Gaps    0;

Qy     338 AGTGGCAGTGACTCTCTTCCTCTTCTCTTCTGCAGCCCCGACACGTCTCTGCTACTGGAC 397
           | | | | || |  |    | ||| ||  |   ||||||||||||||||||||||||
Db    1536 ATTCGAAATGTCAGTGCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGAC 1595

Qy     398 TGTGGTGAGGGCACRTTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTC 457
           |||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db    1596 TGTGGTGAGGGCACATTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTC 1655

Qy     458 CTGGGCACCCTGGCTGCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGG 512
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1656 CTGGGCACCCTGGCTGCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGG 1710


RESULT 10
AAH14250
ID   AAH14250 standard; cDNA; 2976 BP.
XX
AC   AAH14250;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA sequence SEQ ID NO:11557.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;   99JP-0248036.
PR   27-AUG-1999;   99JP-0300253.
PR   11-JAN-2000; 2000JP-0118776.
PR   02-MAY-2000; 2000JP-0183767.
PR   09-JUN-2000; 2000JP-0241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602
PT   full-length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full-length cDNAs -
XX
PS   Claim 8; SEQ ID 11557; 2537pp + CD ROM; English.
XX
CC   The present invention describes primer sets for synthesising 5602
CC   full-length cDNAs defined in the specification. Where a primer set
CC   comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC   to the complementary strand of a polynucleotide which comprises one of
CC   the 5602 nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in
```

```
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC   AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 2976 BP; 712 A; 807 C; 856 G; 601 T; 0 other;

  Query Match             18.2%;  Score 145.8;  DB 22;  Length 2976;
  Best Local Similarity   89.1%;  Pred. No. 9.3e-32;
  Matches  156;  Conservative    1;  Mismatches   18;  Indels    0;  Gaps    0;

Qy     338 AGTGGCAGTGACTCTCTTCCTCTTCTCTTCTGCAGCCCCGACACGTCTCTGCTACTGGAC 397
           | | | | || |  |    | ||| ||  |   ||||||||||||||||||||||||||
Db    1557 ATTCGAAATGTCAGTGCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGAC 1616

Qy     398 TGTGGTGAGGGCACRTTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTC 457
           |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db    1617 TGTGGTGAGGGCACATTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTC 1676

Qy     458 CTGGGCACCCTGGCTGCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGG 512
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1677 CTGGGCACCCTGGCTGCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGG 1731


RESULT 11
ABN59829
ID   ABN59829 standard; cDNA; 2992 BP.
XX
AC   ABN59829;
XX
DT   28-JUN-2002  (first entry)
XX
DE   Novel human coding sequence SEQ ID NO: 240.
XX
KW   Human; antianaemic; vulnerary; antiinflammatory; immunomodulator;
KW   antiinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
KW   neuroprotective; antiparkinsonian; protein therapy; EST;
KW   expressed sequence tag; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200222660-A2.
XX
PD   21-MAR-2002.
XX
PF   10-SEP-2001; 2001WO-US26015.
XX
PR   11-SEP-2000; 2000US-0659671.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Zhou P,  Asundi V,  Zhang J,  Zhao QA,  Ren F;
PI   Xue AJ,  Yang Y,  Wehrman T,  Drmanac RT;
XX
DR   WPI; 2002-292408/33.
DR   P-PSDB; ABB97416.
XX
PT   An isolated polynucleotide for treating diseases associated with its
PT   encoded polypeptide such as cancer and multiple sclerosis -
XX
PS   Claim 1; SEQ ID NO 240; 509pp; English.
XX
CC   The present invention provides the protein and coding sequences of 444
CC   novel human proteins. These were isolated from expressed sequences tags
CC   (ESTs). They can be used to stimulate cell growth, to regulate
CC   haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC   e.g. in burn treatment, to regulate the immune system e.g. to treat
CC   multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC   infertility, to regulate haemostasis or thrombolysis e.g. to treat
CC   stroke and cancer, to screen for drugs, to treat inflammatory conditions
CC   e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
CC   Parkinson's disease. The present sequence is a coding sequence of the
CC   invention.
XX
SQ   Sequence 2992 BP; 725 A; 807 C; 859 G; 601 T; 0 other;

  Query Match             18.2%;  Score 145.8;  DB 24;  Length 2992;
  Best Local Similarity   89.1%;  Pred. No. 9.3e-32;
  Matches  156;  Conservative    1;  Mismatches   18;  Indels    0;  Gaps    0;

Qy     338 AGTGGCAGTGACTCTCTTCCTCTTCTCTTCTGCAGCCCCGACACGTCTCTGCTACTGGAC 397
           | | | | || |  |    | ||| ||  |   ||||||||||||||||||||||||||
Db    1560 ATTCGAAATGTCAGTGCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGAC 1619

Qy     398 TGTGGTGAGGGCACRTTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTC 457
           |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db    1620 TGTGGTGAGGGCACATTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTC 1679

Qy     458 CTGGGCACCCTGGCTGCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGG 512
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1680 CTGGGCACCCTGGCTGCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGG 1734


RESULT 12
AAA52810
ID   AAA52810 standard; cDNA; 2478 BP.
XX
AC   AAA52810;
XX
DT   20-SEP-2000  (first entry)
XX
DE   Human sulphatase G cDNA.
XX
KW   Human sulphatase G; hSG; chromosome 17p11.2; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..2478
FT                   /*tag= a
FT                   /partial
FT                   /product= "hSG"
XX
PN   WO200034327-A1.
XX
PD   15-JUN-2000.
XX
PF   09-DEC-1999;   99WO-AU01092.
XX
PR   09-DEC-1998;   98AU-0007624.
XX
PA   (WOME-) WOMEN'S & CHILDREN'S HOSPITAL.
XX
PI   Hopwood JJ,  Litjens T,  Hu RL;
XX
DR   WPI; 2000-431273/37.
DR   P-PSDB; AAY99850.
XX
PT   Novel isolated DNA sequence which encodes human sulfatase G or its
PT   fragment useful in gene therapy for treating patients suffering from
PT   sulfatase deficiency -
XX
PS   Claim 2; Page 29-30; 33pp; English.
XX
CC   The present sequence encodes human sulphatase G (hSG). hSG is
CC   not a member of the well-characterised CTPSR sulphatase family.
CC   It belongs to a family showing sequence similarity to a sulphatase
CC   from the marine bacterium Alteromonas carrageenovora. The hSG gene
CC   contains 23 exons and is located at chromosome 17p11.2. The present
CC   sequence is clone lambda29.1 from a human testes cDNA library. It was
```

```
CC   isolated using human EST sequences with sequence similarity to the
CC   non-CTPSR family as a probe to screen the library. The cDNA insert was
CC   subcloned and the DNA sequence of both strands was determined. The
CC   sequence may be used to treat a patient suffering from hSG deficiency
CC   by replacing, repairing, or compensating for a DNA sequence within that
CC   patient's genome.
XX
SQ   Sequence 2478 BP; 587 A; 686 C; 709 G; 496 T; 0 other;

  Query Match             18.0%;  Score 144.2;  DB 21;  Length 2478;
  Best Local Similarity   88.6%;  Pred. No. 2.5e-31;
  Matches  155;  Conservative    1;  Mismatches   19;  Indels    0;  Gaps    0;

Qy     338 AGTGGCAGTGACTCTCTTCCTCTTCTCTTCTGCAGCCCCGACACGTCTCTGCTACTGGAC 397
           | | | | || |  |    | ||| ||  |   ||||||||||||||||||||||||||
Db    1486 ATTCGAAATGTCAGTGCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGAC 1545

Qy     398 TGTGGTGAGGGCACRTTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTC 457
           |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db    1546 TGTGGTGAGGGCACGTTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTC 1605

Qy     458 CTGGGCACCCTGGCTGCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGG 512
           |||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    1606 CTGGGCACCCTGGCTACTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGG 1660


RESULT 13
AAS72207
ID   AAS72207 standard; cDNA; 1402 BP.
XX
AC   AAS72207;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #8011.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG08020.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity -
XX
PS   Claim 1; SEQ ID No 8011; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1402 BP; 338 A; 371 C; 377 G; 316 T; 0 other;

  Query Match             17.7%;  Score 141.6;  DB 23;  Length 1402;
  Best Local Similarity   99.3%;  Pred. No. 1.1e-30;
  Matches  141;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     371 AGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACRTTTGGGCAGCTGTGCCGT 430
           |||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db     674 AGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACATTTGGGCAGCTGTGCCGT 733

Qy     431 CATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCTGCTGTGTTTGTGTCCCAC 490
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     734 CATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCTGCTGTGTTTGTGTCCCAC 793

Qy     491 CTGCACGCAGATCACCACACGG 512
           |||||||||||||||||||||
Db     794 CTGCACGCAGATCACCACACGG 815


RESULT 14
AAS98934
ID   AAS98934 standard; DNA; 139 BP.
XX
AC   AAS98934;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Human prostate cancer predisposing gene (HPC2) DNA partial exon #17.
XX
KW   Human; mouse; HPC2; prostate cancer; neoplastic growth; cytostatic; ds;
KW   gene therapy; prostate cancer predisposing gene.
XX
OS   Homo sapiens.
XX
PN   WO200185911-A2.
XX
PD   15-NOV-2001.
XX
PF   07-MAY-2001; 2001WO-US14602.
XX
PR   05-MAY-2000; 2000US-0564805.
XX
PA   (MYRI-) MYRIAD GENETICS INC.
PA   (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI   Tavtigian SV,  Teng DHF,  Simard J,  Rommens JM;
XX
DR   WPI; 2002-066599/09.
XX
PT   Novel nucleic acid sequence encoding HPC2 polypeptide, which is marker
PT   for prostate cancer, is useful in gene therapy techniques to restore
PT   HPC2 normal levels by which neoplastic growth is suppressed in
PT   recipient cell -
XX
PS   Claim 9; Page 140; 239pp; English.
XX
CC   The invention relates to a human prostate cancer predisposing gene coding
CC   for an HPC2 polypeptide. The DNA and protein sequences are useful as
CC   diagnostic reagents for identifying a mutant HPC2 nucleotide sequence in
CC   a suspected mutant HPC2 allele by comparing the sequence of the suspected
```

```
CC   mutant HPC2 allele with a wild-type HPC2 sequence. The sequences are also
CC   useful for detecting an alteration in HPC2, where the alteration is
CC   associated with cancer in a human. The method involves analysing an HPC2
CC   gene or an HPC2 gene expression product from a tissue of the human. The
CC   HPC2 gene is useful as a marker for prostate cancer and can be used in
CC   gene therapy techniques to suppress neoplastic growth of recipient cells
CC   which carry the mutant HPC2 allele. The sequences represent DNA encoding
CC   human and mouse HPC2 and fragments of HPC2.
XX
SQ   Sequence 139 BP; 23 A; 43 C; 43 G; 30 T; 0 other;

  Query Match             17.3%;  Score 138.6;  DB 24;  Length 139;
  Best Local Similarity   99.3%;  Pred. No. 2.4e-30;
  Matches  138;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     373 CCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACRTTTGGGCAGCTGTGCCGTCA 432
           |||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db       1 CCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACATTTGGGCAGCTGTGCCGTCA 60

Qy     433 TTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCTGCTGTGTTTGTGTCCCACCT 492
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCTGCTGTGTTTGTGTCCCACCT 120

Qy     493 GCACGCAGATCACCACACG 511
           ||||||||||||||||||
Db     121 GCACGCAGATCACCACACG 139


RESULT 15
AAS99131
ID   AAS99131 standard; cDNA; 2470 BP.
XX
AC   AAS99131;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Mouse ELAC2 cDNA.
XX
KW   Human; mouse; HPC2; prostate cancer; neoplastic growth; cytostatic; ss;
KW   gene therapy; prostate cancer predisposing gene; chimpanzee; gorilla;
KW   sequencing primer; PCR primer.
XX
OS   Mus musculus.
XX
PN   WO200185911-A2.
XX
PD   15-NOV-2001.
XX
PF   07-MAY-2001; 2001WO-US14602.
XX
PR   05-MAY-2000; 2000US-0564805.
XX
PA   (MYRI-) MYRIAD GENETICS INC.
PA   (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI   Tavtigian SV,  Teng DHF,  Simard J,  Rommens JM;
XX
DR   WPI; 2002-066599/09.
DR   P-PSDB; AAU73591.
XX
PT   Novel nucleic acid sequence encoding HPC2 polypeptide, which is marker
PT   for prostate cancer, is useful in gene therapy techniques to restore
PT   HPC2 normal levels by which neoplastic growth is suppressed in
PT   recipient cell  -
XX
PS   Claim 82; Page 192-195; 239pp; English.
XX
CC   The invention relates to a human prostate cancer predisposing gene coding
CC   for an HPC2 polypeptide. The DNA and protein sequences are useful as
CC   diagnostic reagents for identifying a mutant HPC2 nucleotide sequence in
CC   a suspected mutant HPC2 allele by comparing the sequence of the suspected
CC   mutant HPC2 allele with a wild-type HPC2 sequence. The sequences are also
CC   useful for detecting an alteration in HPC2, where the alteration is
CC   associated with cancer in a human. The method involves analysing an HPC2
CC   gene or an HPC2 gene expression product from a tissue of the human. The
CC   HPC2 gene is useful as a marker for prostate cancer and can be used in
CC   gene therapy techniques to suppress neoplastic growth of recipient cells
CC   which carry the mutant HPC2 allele. The sequences represent primers used
CC   in the methods of the invention, cDNA encoding human and mouse HPC2 and
CC   cDNA encoding HPC2 paralogues and orthologues.
XX
SQ   Sequence 2470 BP; 614 A; 664 C; 679 G; 510 T; 3 other;

  Query Match             12.8%;  Score 102.2;  DB 24;  Length 2470;
  Best Local Similarity   74.9%;  Pred. No. 4.2e-19;
  Matches  128;  Conservative    0;  Mismatches   43;  Indels    0;  Gaps    0;

Qy     342 GCAGTGACTCTCTTCCTCTTCTCTTCTGCAGCCCCGACACGTCTCTGCTACTGGACTGTG 401
           | | || |  |    | ||  ||  |   ||||| |||| ||| |||| ||||| ||||
Db    1451 GAAATGTCAGTTCCACACTCGTCAACCTAAGCCCTGACAAGTCAGTGCTCCTGGATTGTG 1510

Qy     402 GTGAGGGCACRTTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGG 461
           | || ||||| ||||||||| |||||||||||||||| | ||  | ||| | ||| |
Db    1511 GAGAAGGCACTTTTGGGCAGTTGTGCCGTCATTACGGACAGCAAATAGACCGAGTCTTAT 1570

Qy     462 GCACCCTGGCTGCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGG 512
           ||| |||  | |||||||||||||||||||||||| || ||||||||||
Db    1571 GCAGCCTCACGGCTGTGTTTGTGTCCCACCTGCACGCCGACCACCACACGG 1621


Search completed: May 16, 2003, 03:42:22
Job time : 263.416 secs
```

GenCore version 5.1.5

OM nucleic - nucleic search, using sw model

Run on: May 16, 2003, 01:11:33 ; Search time 111.371 Seconds
(without alignments)
10110.387 Million cell updates/sec

Title: US-09-434-382-28_COPY_1_500
Perfect score: 500
Sequence: 1 tatcaggtgactgaattcta.....aagctctgagggactgacgt 500

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	Query Match %	Length	DB	ID	Description
	1	500	100.0	26664	21	AAA60207	Human prostate can
	2	500	100.0	26664	24	AAS98942	Human prostate can
	3	120.4	24.1	7680	21	AAX89439	14-3-3 sigma trans
c	4	120.2	24.0	10708	22	AAK66570	Human immune/haema
	5	120	24.0	1503900	22	AAK95240	Human neuregulin-1
	6	120	24.0	1503900	22	AAK96733	Human neuregulin-1
	7	119.8	24.0	4045	22	AAL06030	Human reproductive
	8	119.8	24.0	4045	22	AAL06031	Human reproductive
	9	119.8	24.0	4045	23	ABL98595	Human testicular a
	10	119.8	24.0	4045	23	ABL98596	Human testicular a
c	11	119.8	24.0	33780	22	AAH24652	Nucleotide sequenc
c	12	119.6	23.9	1160	16	AAQ85372	RANTES cDNA sequen
c	13	119.6	23.9	1160	19	AAV36275	cDNA containing th
c	14	119.6	23.9	1160	21	AAF21050	Human low adenosin
c	15	119.6	23.9	1160	21	AAA74884	Human chemokine co
c	16	119.6	23.9	1160	21	AAA34928	Human adenosine re
c	17	119.6	23.9	1160	24	ABK64521	Human benign prost
c	18	119.6	23.9	1160	24	ABL62876	Breast cancer rela
c	19	119.6	23.9	2176	21	AAF21051	Human low adenosin
c	20	119.6	23.9	2176	21	AAA34929	Human adenosine re
	21	119.6	23.9	30626	22	AAK67051	Human immune/haema
	22	118.6	23.7	49999	20	AAZ23902	Human LOBO homolog
c	23	118.4	23.7	465237	24	ABQ87681	Human oestrogen re
c	24	118.4	23.7	465237	24	ABA90193	Human oestrogen re
	25	118.2	23.6	5159	22	AAL37342	Human musculoskele
c	26	118.2	23.6	5441	22	AAI98937	Human excretory re
c	27	118.2	23.6	5441	22	AAI63287	Human kidney relat
	28	118.2	23.6	6834	22	AAL05314	Human reproductive
c	29	118.2	23.6	6834	22	AAS28901	Human immunoglobul
c	30	118	23.6	1115	22	AAS40046	Genomic sequence #
c	31	118	23.6	1115	22	AAS40047	Genomic sequence #
c	32	118	23.6	1115	22	AAK91463	Human digestive sy
c	33	118	23.6	1115	22	AAK91464	Human digestive sy
	34	118	23.6	6428	22	AAK78562	Human immune/haema
	35	117.8	23.6	727	22	AAH04798	Human cDNA clone (
	36	117.8	23.6	2869	22	AAH17472	Human cDNA sequenc
	37	117.6	23.5	6186	22	AAS28641	Genomic sequence #
	38	117.6	23.5	6906	22	AAK66219	Human immune/haema
c	39	117.4	23.5	6618	22	ABA18101	Human nervous syst
c	40	117.4	23.5	12267	22	AAK85733	Human immune/haema
c	41	117.2	23.4	433	22	AAK96605	Human neuregulin g
c	42	117.2	23.4	433	22	AAK98098	Human neuregulin g
	43	117.2	23.4	582	22	AAH10027	Human cDNA clone (
c	44	117.2	23.4	5796	22	AAS42029	Genomic sequence #
c	45	117.2	23.4	18010	22	AAK67807	Human immune/haema

ALIGNMENTS

RESULT 1
AAA60207
ID AAA60207 standard; DNA; 26664 BP.
XX
AC AAA60207;
XX
DT 07-DEC-2000 (first entry)
XX
DE Human prostate cancer predisposing gene HPC2 genomic sequence.
XX
KW Human; prostate cancer predisposing gene; HPC2; chromosome 17p;
KW gene therapy; peptide therapy; drug design; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 910..26039
FT /*tag= a
FT /product= "HPC2"
FT /note= "this sequence contains introns"
FT /transl_except= (pos:23892..23895,aa:Glu)
FT exon 910..1154
FT /*tag= b
FT /number= 1
FT exon 1736..1786
FT /*tag= c
FT /number= 2
FT exon 1925..1995
FT /*tag= d
FT /number= 3
FT exon 3025..3089
FT /*tag= e

```
FT                   /number= 4
FT   exon            4361..4418
FT                   /*tag=  f
FT                   /number= 5
FT   exon            5582..5650
FT                   /*tag=  g
FT                   /number= 6
FT   exon            7075..7194
FT                   /*tag=  h
FT                   /number= 7
FT   exon            8186..8244
FT                   /*tag=  i
FT                   /number= 8
FT   exon            12878..12936
FT                   /*tag=  j
FT                   /number= 9
FT   exon            13032..13104
FT                   /*tag=  k
FT                   /number= 10
FT   exon            13756..13868
FT                   /*tag=  l
FT                   /number= 11
FT   exon            15283..15378
FT                   /*tag=  m
FT                   /number= 12
FT   exon            16278..16416
FT                   /*tag=  n
FT                   /number= 13
FT   exon            16498..16583
FT                   /*tag=  o
FT                   /number= 14
FT   exon            18583..18701
FT                   /*tag=  p
FT                   /number= 15
FT   exon            20349..20445
FT                   /*tag=  q
FT                   /number= 16
FT   exon            22172..22310
FT                   /*tag=  r
FT                   /number= 17
FT   exon            22879..22917
FT                   /*tag=  s
FT                   /number= 18
FT   exon            23045..23154
FT                   /*tag=  t
FT                   /number= 19
FT   exon            23795..23895
FT                   /*tag=  u
FT                   /number= 20
FT   exon            23973..24093
FT                   /*tag=  v
FT                   /number= 21
FT   exon            24354..24432
FT                   /*tag=  w
FT                   /number= 22
FT   exon            25026..25170
FT                   /*tag=  x
FT                   /number= 23
FT   exon            25812..26036
FT                   /*tag=  y
FT                   /number= 24
FT   polyA_signal    26447..26452
FT                   /*tag=  z
XX
PN   WO200027864-A1.
XX
PD   18-MAY-2000.
XX
PF   05-NOV-1999;   99WO-US26055.
XX
PR   06-NOV-1998;   98US-0107468.
XX
PA   (MYRI-) MYRIAD GENETICS INC.
XX
PI   Tavtigian SV,  Teng DHF,  Simard J,  Rommens JM;
XX
DR   WPI; 2000-376481/32.
DR   P-PSDB; AAB07228.
XX
PT   Human prostate cancer (HPC)2 nucleic acids, polypeptides, and
PT   antibodies, useful for treatment and diagnosis of prostate cancer  -
XX
PS   Claim 3; Page 108-122; 157pp; English.
XX
CC   The present sequence is the genomic sequence of the human prostate
CC   cancer predisposing gene HPC2, which is found on chromosome 17p. Some
CC   alleles of this gene cause a predisposition to cancer, particularly
CC   prostate cancer. This gene and its protein can be used in peptide and
CC   gene therapy for cancer patients, as well as being useful as diagnostic
CC   tools (both for cancer sufferers and those with a predisposition to the
CC   disease) and in the production of cancer drugs. This sequence was
CC   isolated by cloning and sequencing the region of the genome which
CC   appeared to cause a predisposition to prostate cancer.
XX
SQ   Sequence 26664 BP; 6173 A; 6300 C; 6519 G; 7661 T; 11 other;

  Query Match             100.0%;  Score 500;  DB 21;  Length 26664;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-125;
  Matches  500;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TATCAGGTGACTGAATTCTATATTCTGAAGTAGGAGATACTGTTATTGCTGTTATTACAT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 TATCAGGTGACTGAATTCTATATTCTGAAGTAGGAGATACTGTTATTGCTGTTATTACAT 60

Qy      61 TTTACACATAAGAAAGCTGAGGCTCTGAGAGGTCAAGATCACGCAGCTAACAAATGAGCC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TTTACACATAAGAAAGCTGAGGCTCTGAGAGGTCAAGATCACGCAGCTAACAAATGAGCC 120

Qy     121 AAGACTCTTGCTTTAGAGCTTGTCCTCTATTCTTGCTTTTCTTTCCAAAAAACACTACAA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AAGACTCTTGCTTTAGAGCTTGTCCTCTATTCTTGCTTTTCTTTCCAAAAAACACTACAA 180

Qy     181 TTTTTGTTTTGTTTTGTTTTGTTTTGAGACAGGGTCTCGAGGTGTCACCCAGGCTGGAGT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TTTTTGTTTTGTTTTGTTTTGTTTTGAGACAGGGTCTCGAGGTGTCACCCAGGCTGGAGT 240

Qy     241 GCAGTGGCGCGATTTCGACTCACCGCAACCTCCGCCTCCGCGCTTAAGCGATTCTCCTGC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GCAGTGGCGCGATTTCGACTCACCGCAACCTCCGCCTCCGCGCTTAAGCGATTCTCCTGC 300

Qy     301 CTCAGCCTCCCAAGTAGCTGGGACTACAAGCTCGGGACACCACGTAAAAATGATCAAGTT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CTCAGCCTCCCAAGTAGCTGGGACTACAAGCTCGGGACACCACGTAAAAATGATCAAGTT 360

Qy     361 CTAACATGTATGCATACGAATTACAATGGAAATAAAATTAGCAAAGCGCTTATGCTAATG 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CTAACATGTATGCATACGAATTACAATGGAAATAAAATTAGCAAAGCGCTTATGCTAATG 420

Qy     421 CTCAATACAATTGATTTCCTCACATTTAATCCTCACAACCACTACAACCACCTCTAACTC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CTCAATACAATTGATTTCCTCACATTTAATCCTCACAACCACTACAACCACCTCTAACTC 480

Qy     481 AAGCTCTGAGGGACTGACGT 500
           ||||||||||||||||||||
Db     481 AAGCTCTGAGGGACTGACGT 500


RESULT 2
AAS98942
ID   AAS98942 standard; DNA; 26664 BP.
XX
AC   AAS98942;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Human prostate cancer predisposing gene (HPC2) genomic DNA.
```

```
XX
KW   Human; mouse; HPC2; prostate cancer; neoplastic growth; cytostatic; ds;
KW   gene therapy; prostate cancer predisposing gene.
XX
OS   Homo sapiens.
XX
PN   WO200185911-A2.
XX
PD   15-NOV-2001.
XX
PF   07-MAY-2001; 2001WO-US14602.
XX
PR   05-MAY-2000; 2000US-0564805.
XX
PA   (MYRI-) MYRIAD GENETICS INC.
PA   (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI   Tavtigian SV,  Teng DHF,  Simard J,  Rommens JM;
XX
DR   WPI; 2002-066599/09.
XX
PT   Novel nucleic acid sequence encoding HPC2 polypeptide, which is marker
PT   for prostate cancer, is useful in gene therapy techniques to restore
PT   HPC2 normal levels by which neoplastic growth is suppressed in
PT   recipient cell
XX
PS   Claim 3; Page 143-157; 239pp; English.
XX
CC   The invention relates to a human prostate cancer predisposing gene coding
CC   for an HPC2 polypeptide. The DNA and protein sequences are useful as
CC   diagnostic reagents for identifying a mutant HPC2 nucleotide sequence in
CC   a suspected mutant HPC2 allele by comparing the sequence of the suspected
CC   mutant HPC2 allele with a wild-type HPC2 sequence. The sequences are also
CC   useful for detecting an alteration in HPC2, where the alteration is
CC   associated with cancer in a human. The method involves analysing an HPC2
CC   gene or an HPC2 gene expression product from a tissue of the human. The
CC   HPC2 gene is useful as a marker for prostate cancer and can be used in
CC   gene therapy techniques to suppress neoplastic growth of recipient cells
CC   which carry the mutant HPC2 allele. The sequences represent DNA encoding
CC   human and mouse HPC2 and fragments of HPC2.
XX
SQ   Sequence 26664 BP; 6173 A; 6300 C; 6519 G; 7661 T; 11 other;

  Query Match             100.0%;  Score 500;  DB 24;  Length 26664;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-125;
  Matches  500;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TATCAGGTGACTGAATTCTATATTCTGAAGTAGGAGATACTGTTATTGCTGTTATTACAT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 TATCAGGTGACTGAATTCTATATTCTGAAGTAGGAGATACTGTTATTGCTGTTATTACAT 60

Qy      61 TTTACACATAAGAAAGCTGAGGCTCTGAGAGGTCAAGATCACGCAGCTAACAAATGAGCC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TTTACACATAAGAAAGCTGAGGCTCTGAGAGGTCAAGATCACGCAGCTAACAAATGAGCC 120

Qy     121 AAGACTCTTGCTTTAGAGCTTGTCCTCTATTCTTGCTTTTCTTTCCAAAAAACACTACAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AAGACTCTTGCTTTAGAGCTTGTCCTCTATTCTTGCTTTTCTTTCCAAAAAACACTACAA 180

Qy     181 TTTTTGTTTTGTTTTGTTTTGTTTTGAGACAGGGTCTCGAGGTGTCACCCAGGCTGGAGT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TTTTTGTTTTGTTTTGTTTTGTTTTGAGACAGGGTCTCGAGGTGTCACCCAGGCTGGAGT 240

Qy     241 GCAGTGGCGCGATTTCGACTCACCGCAACCTCCGCCTCCGCGCTTAAGCGATTCTCCTGC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GCAGTGGCGCGATTTCGACTCACCGCAACCTCCGCCTCCGCGCTTAAGCGATTCTCCTGC 300

Qy     301 CTCAGCCTCCCAAGTAGCTGGGACTACAAGCTCGGGACACCACGTAAAAATGATCAAGTT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CTCAGCCTCCCAAGTAGCTGGGACTACAAGCTCGGGACACCACGTAAAAATGATCAAGTT 360

Qy     361 CTAACATGTATGCATACGAATTACAATGGAAATAAAATTAGCAAAGCGCTTATGCTAATG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CTAACATGTATGCATACGAATTACAATGGAAATAAAATTAGCAAAGCGCTTATGCTAATG 420

Qy     421 CTCAATACAATTGATTTCCTCACATTTAATCCTCACAACCACTACAACCACCTCTAACTC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CTCAATACAATTGATTTCCTCACATTTAATCCTCACAACCACTACAACCACCTCTAACTC 480

Qy     481 AAGCTCTGAGGGACTGACGT 500
          ||||||||||||||||||||
Db     481 AAGCTCTGAGGGACTGACGT 500


RESULT 3
AAX89439
ID   AAX89439 standard; DNA; 7680 BP.
XX
AC   AAX89439;
XX
DT   15-FEB-2000  (first entry)
XX
DE   14-3-3 sigma transcription regulatory region.
XX
KW   14-3-3 sigma; HME1; stratifin; p53; diagnosis; cancer; psoriasis; polyp;
KW   psoriasis; wart; inflammatory disease; proliferation; ss;
KW   transcription regulatory region.
XX
OS   Homo sapiens.
XX
PN   WO9931240-A2.
XX
PD   24-JUN-1999.
XX
PF   18-DEC-1998;   98WO-US26924.
XX
PR   18-DEC-1997;   97US-0069416.
PR   15-DEC-1998;   98US-0210748.
XX
PA   (UYJO ) UNIV JOHNS HOPKINS.
XX
PI   Hermeking H,  Vogelstein B,  Kinzler KW;
XX
DR   WPI; 2000-022907/02.
XX
PT   Use of 14-3-3 sigma polypeptides and nucleic acids for the diagnosis or
PT   treatment of cancer
XX
PS   Claim 52; Page 68-71; 73pp; English.
XX
CC   This is the 14-3-3 sigma transcriptional regulatory region nucleotide
CC   sequence. 14-3-3 sigma is a member of the 14-3-3 protein family and is
CC   also known as HME1 or stratifin. 14-3-3 sigma expression is regulated by
CC   p53 and exogenous expression of 14-3-3 sigma results in G2 block. The
CC   14-3-3 sigma nucleotide and amino acid sequences are used in the
CC   invention to develop agents for the diagnosis, susceptibility
CC   determination and treatment of cancer. The amino acid sequence can be
CC   used in method for suppressing the growth of tumour cells. The 14-3-3
CC   sigma polypeptides can mediate cell cycle arrest upon damage to cellular
CC   DNA. 14-3-3 sigma probes can be used for diagnosing, testing
CC   susceptibility to or treating cancers and identifying agents for treating
CC   cancers. They can also be used to treat other proliferative diseases,
CC   e.g. psoriasis, polyps, warts, and inflammatory diseases. The 14-3-3
CC   sigma antisense oligonucleotides can be used for promoting the
CC   proliferation and growth of cells.
XX
SQ   Sequence 7680 BP; 1798 A; 2053 C; 2000 G; 1829 T; 0 other;

  Query Match             24.1%;  Score 120.4;  DB 21;  Length 7680;
  Best Local Similarity   74.3%;  Pred. No. 1.5e-22;
  Matches  165;  Conservative    0;  Mismatches   56;  Indels    1;  Gaps    1;

Qy     124 ACTCTTGCTTTAGAGCTTGTCCTCTATTCTTGCTTTTCTTTCCAAAAAACACTACAATTT 183
           | |||   |  | || |  |  ||||||||   | |  |   |||| || ||
Db    5886 AGTCTGATTCCAAAGATACTAGTCTATTCTGTATCTCATAGACAAACAATACATATTCAC 5945
```

```
Qy     184 TTGTTTTGTTTTGTTTTGTTTTGAGACAGGGTCTCGAGGTGTCACCCAGGCTGGAGTGCA 243
           || |||  | ||||||||||||||||| | |||| |   ||||||||||||||||||||
Db    5946 TTTTTTGTTGTTGTTTTGTTTTGAGACGGAGTCTTGCTCTGTCACCCAGGCTGGAGTGCA 6005

Qy     244 GTGGCGCGATTTCGACTCACCGCAACCTCCGCCT-CCGCGCTTAAGCGATTCTCCTGCCT 302
           ||||||| || ||| ||||| ||||| ||||||| ||| | | ||||||||||||||||||
Db    6006 GTGGCGCCATCTCGGCTCACTGCAACGTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCT 6065

Qy     303 CAGCCTCCCAAGTAGCTGGGACTACAAGCTCGGGACACCACG 344
           ||||||||| ||||||||||||||.|| ||  | | ||||| |
Db    6066 CAGCCTCCCGAGTAGCTGGGACTACAGGCATGTGCCACCATG 6107
```

RESULT 4
AAK66570/c
ID AAK66570 standard; DNA; 10708 BP.
XX
AC AAK66570;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:21382.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.

```
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-483426/52.
XX
PT   Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating cancers and
PT   metastasis -
XX
PS   Disclosure; SEQ ID NO 21382; 3071pp + Sequence Listing; English.
XX
CC   AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC   amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC   activity, and can be used in gene therapy and vaccine production. (I)
CC   proteins and polynucleotides may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate (I) expression. For
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
CC   supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting
CC   the nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention.
XX
SQ   Sequence 10708 BP; 2285 A; 3103 C; 3106 G; 2214 T; 0 other;

  Query Match             24.0%;  Score 120.2;  DB 22;  Length 10708;
  Best Local Similarity   83.0%;  Pred. No. 1.9e-22;
  Matches  137;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy     180 ATTTTTGTTTTGTTTTGTTTTGTTTTGAGACAGGGTCTCGAGGTGTCACCCAGGCTGGAG 239
           | |||| |||| |||| |||| ||||||| | | ||||||   |||| ||||||||||||
Db    7429 ACTTTTTTTTTTTTTTTTTTTTTTTGAGTCGGAGTCTCGCTCTGTCGCCCAGGCTGGAG 7370

Qy     240 TGCAGTGGCGCGATTTCGACTCACCGCAACCTCCGCCTCCGCGCTTAAGCGATTCTCCTG 299
           |||||||| ||||| ||| ||||| |||| |||||||||||| | | | || |||||||||
Db    7369 TGCAGTGGTGCGATCTCGGCTCACTGCAAGCTCCGCCTCCGGGGTCACGCCATTCTCCTG 7310

Qy     300 CCTCAGCCTCCCAAGTAGCTGGGACTACAAGCTCGGGACACCACG 344
           ||||||||||||||||||||||||| ||| || |  | |||||||
Db    7309 CCTCAGCCTCCCAAGTAGCTGGGACAACAGGCGCCCGCCACCACG 7265
```

```
RESULT 5
AAK95240
ID   AAK95240 standard; DNA; 1503900 BP.
XX
AC   AAK95240;
XX
DT   17-DEC-2001  (first entry)
XX
DE   Human neuregulin-1 gene.
XX
KW   Human; neuregulin-1 associated gene 1; NRG1AG1; Schizophrenia gene;
KW   gene therapy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200164876-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-US06376.
XX
PR   28-FEB-2000; 2000US-0515715.
XX
PA   (DECO-) DECODE GENETICS EHF.
XX
PI   Stefansson H,  Steinthorsdottir V,  Gulcher JR;
XX
DR   WPI; 2001-550179/61.
DR   P-PSDB; AAG67900, AAG67901, AAG67902, AAG67903, AAG67904, AAG67905,
DR   AAG67906, AAG67907, AAG67908, AAG67909, AAG67910, AAG67911, AAG67912,
DR   AAG67913, AAG67914, AAG67915, AAG67916, AAG67917, AAG67918, AAG67919,
DR   AAG67920, AAG67921, AAG67922, AAG67923, AAG67924, AAG67925, AAG67926,
DR   AAG67927, AAG67928, AAG67929, AAG67930, AAG67931, AAG67932, AAG67933,
DR   AAG67934, AAG67935, AAG67936, AAG67937.
XX
PT   Neuregulin-1 associated gene 1 nucleic acids and fragments, useful for
PT   preventing diagnosing and treating schizophrenia -
XX
PS   Disclosure; Page 90-501; 750pp; English.
XX
CC   This sequence represents the human neuregulin-1 associated gene 1
CC   (NRG1AG1) of the invention. The NRG1AG1 gene is also referred to as the
CC   human Schizophrenia gene. The invention also relates to fragments or
CC   variants of the gene and the NRG1AG1 polypeptides they encode. The
CC   NRG1AG1 nucleic acids and polypeptides may be used in the prevention,
CC   diagnosis and treatment of diseases associated with inappropriate NRG1AG1
CC   expression. For example, they may be used to treat disorders associated
CC   with decreased expression by rectifying mutations or deletions in a
CC   patient's genome that affect the activity of NRG1AG1 by expressing
CC   inactive proteins or to supplement the patients own production of
CC   NRG1AG1. Additionally, the gene may be used to produce NRG1AG1
CC   polypeptides, by inserting the nucleic acids into a host cell and
CC   culturing the cell to express the protein. The gene may also be used as
CC   DNA probes and primers in diagnostic assays to detect and quantitate the
CC   presence of similar nucleic acids in samples, and therefore which
CC   patients may be in need of restorative therapy. The NRG1AG1 polypeptides
CC   may also be used as antigens in the production of antibodies against
CC   NRG1AG1 and in assays to identify modulators of NRG1AG1 expression and
CC   activity. Anti-NRG1AG1 antibodies and antagonists may also be used to
CC   down regulate expression and activity. Anti-NRG1AG1 antibodies may
CC   also be used as diagnostic agents for detecting the presence of NRG1AG1
CC   polypeptides in samples. NRG1AG1 Is associated with schizophrenia which
CC   may be prevented, diagnosed and/or treated by the above methods.
XX
SQ   Sequence 1503900 BP; 452487 A; 281874 C; 288074 G; 480092 T; 1373 other;

  Query Match             24.0%;  Score 120;  DB 22;  Length 1503900;
  Best Local Similarity   82.8%;  Pred. No. 1.1e-21;
  Matches  149;  Conservative    0;  Mismatches   30;  Indels    1;  Gaps    1;

Qy     166 CAAAAAACACTACAATTTTTGTTTTGTTTTGTTTTGTTTTGAGACAGGGTCTCGAGGTGT 225
           || | || |||||  ||||| |||| |||| |||| ||||||||| | ||||||   |||
```

```
Db 532365 CAGAGAAAACTACTTTTTTTTTTTTTTTTTTTTTTTTTTTGAGACGGAGTCTCGCTCTGT 532424

Qy    226 CACCCAGGCTGGAGTGCAGTGGCGCGATTTCGACTCACCGCAACCTCCGCCTCCGCGCTT 285
          | ||||||||||||||||||||||| ||| ||| ||||| |||| |||||||||||  | ||
Db 532425 CGCCCAGGCTGGAGTGCAGTGGCGGGATCTCGGCTCACTGCAAGCTCCGCCTCCGGGGTT 532484

Qy    286 -AAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACTACAAGCTCGGGACACCACG 344
           | || ||||||||||||||||||||||||||||||||||||||||| || |  | ||| |||
Db 532485 CACGCCATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACTACAGGCGCCCGCCACTACG 532544


RESULT 6
AAK96733
ID   AAK96733 standard; DNA; 1503900 BP.
XX
AC   AAK96733;
XX
DT   17-DEC-2001  (first entry)
XX
DE   Human neuregulin-1 gene.
XX
KW   Human; neuregulin 1 gene; schizophrenia; gene therapy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200164877-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-US06377.
XX
PR   28-FEB-2000; 2000US-0515716.
XX
PA   (DECO-) DECODE GENETICS EHF.
XX
PI   Stefansson H,  Steinthorsdottir V,  Gulcher JR;
XX
DR   WPI; 2001-514841/56.
DR   P-PSDB; AAG67938, AAG67939, AAG67940, AAG67941, AAG67942, AAG67943,
DR   AAG67944, AAG67945, AAG67946, AAG67947, AAG67948, AAG67949,
DR   AAG67950, AAG67951, AAG67952, AAG67953, AAG67954, AAG67955,
DR   AAG67956, AAG67957, AAG67958, AAG67959, AAG67960, AAG67961,
DR   AAG67962, AAG67963, AAG67964, AAG67965, AAG67966, AAG67967,
DR   AAG67968, AAG67969, AAG67970, AAG67971, AAG67972, AAG67973,
DR   AAG67974, AAG67975.
XX
PT   Neuregulin 1 nucleic acids and proteins useful for diagnosing
PT   preventing and treating schizophrenia -
XX
PS   Disclosure; Page 345-756; 756pp; English.
XX
CC   This sequence represents the human neuregulin 1 gene of the invention.
CC   The invention also relates to fragments or variants of the neuregulin 1
CC   gene. The gene and its proteins may be used in the prevention, diagnosis
CC   and treatment of diseases associated with inappropriate neuregulin 1
CC   expression, such as schizophrenia. For example they may be used to treat
CC   disorders associated with decreased neuregulin 1 expression by rectifying
CC   mutations or deletions in a patient's genome that affect the activity of
CC   neuregulin 1 by expressing inactive proteins or to supplement the
CC   patients own production of polypeptides. Additionally, the gene may be
CC   used to produce the neuregulin 1 protein, by inserting the nucleic acids
CC   into a host cell and culturing the cell to express the protein. The gene
CC   and its complementary sequences may also be used as DNA probes in
CC   diagnostic assays to detect and quantitate the presence of similar
CC   nucleic acids in samples, and therefore which patients may be in need of
CC   restorative therapy. The protein may also be used as antigens in the
CC   production of antibodies against neuregulin 1 and in assays to identify
CC   modulators of neuregulin 1 expression and activity. The antibodies and
CC   antagonists may also be used to down regulate expression and activity.
CC   The antibodies may also be used as diagnostic agents for detecting the
CC   presence of neuregulin 1 in samples.
XX
SQ   Sequence 1503900 BP; 452487 A; 281874 C; 288074 G; 480092 T; 1373 other;

  Query Match             24.0%;  Score 120;  DB 22;  Length 1503900;
  Best Local Similarity   82.8%;  Pred. No. 1.1e-21;
  Matches  149;  Conservative    0;  Mismatches   30;  Indels    1;  Gaps    1;

Qy    166 CAAAAAACACTACAATTTTTGTTTTGTTTTGTTTTGTTTTGAGACAGGGTCTCGAGGTGT 225
          || | || |||||  ||||| |||| |||| |||| |||||||||| | ||||||   |||
Db 532365 CAGAGAAAACTACTTTTTTTTTTTTTTTTTTTTTTTTTTTGAGACGGAGTCTCGCTCTGT 532424

Qy    226 CACCCAGGCTGGAGTGCAGTGGCGCGATTTCGACTCACCGCAACCTCCGCCTCCGCGCTT 285
          | ||||||||||||||||||||||| ||| ||| ||||| |||| |||||||||||  | ||
Db 532425 CGCCCAGGCTGGAGTGCAGTGGCGGGATCTCGGCTCACTGCAAGCTCCGCCTCCGGGGTT 532484

Qy    286 -AAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACTACAAGCTCGGGACACCACG 344
           | || ||||||||||||||||||||||||||||||||||||||||| || |  | ||| |||
Db 532485 CACGCCATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACTACAGGCGCCCGCCACTACG 532544


RESULT 7
AAL06030
ID   AAL06030 standard; DNA; 4045 BP.
XX
AC   AAL06030;
XX
DT   21-NOV-2001  (first entry)
XX
DE   Human reproductive system related antigen DNA SEQ ID NO: 8718.
XX
KW   Human; reproductive system related antigen; reproductive system disorder;
KW   cancer; gene therapy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200155320-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01339.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
```

PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465570/50.
XX
PT Isolated nucleic acid molecule encoding a reproductive system antigen
PT is used in preventing, treating or ameliorating a medical condition -
XX
PS Disclosure; SEQ ID NO 8718; 1297pp + Sequence Listing; English.
XX
CC The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a genomic sequence encoding a
CC protein of the invention.
XX
SQ Sequence 4045 BP; 1044 A; 994 C; 1001 G; 1006 T; 0 other;

Query Match 24.0%; Score 119.8; DB 22; Length 4045;
Best Local Similarity 73.0%; Pred. No. 1.8e-22;
Matches 154; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

```
Qy    133 TTAGAGCTTGTCCTCTATTCTTGCTTTTCTTTCCAAAAAACACTACAATTTTTGTTTTGT 192
          ||  |  ||| | ||| |         ||| ||||        || |  | ||| | |
Db   1603 TTCTACATTGGCTTCTTTCTCGTTCCTTCCTTCCTTCCCTTCCTTCCTTCTTTCTCTCTC 1662

Qy    193 TTTGTTTTGTTTTGAGACAGGGTCTCGAGGTGTCACCCAGGCTGGAGTGCAGTGGCGCGA 252
          | | |||  ||||||||||| |||||    |||| ||||||||||||||||||||||| |||
Db   1663 TCTCTTTCTTTTTGAGACAGAGTCTCACTCTGTCGCCCAGGCTGGAGTGCAGTGGCACGA 1722

Qy    253 TTTCGACTCACCGCAACCTCCGCCTCCGCGCTTAAGCGATTCTCCTGCCTCAGCCTCCCA 312
          | ||  ||||| |||||||| ||||| | | | ||||||||||||||||||||||||||||
Db   1723 TCTCAGCTCACTGCAACCTCTGCCTCTGAGTTCAAGCGATTCTCCTGCCTCAGCCTCCCA 1782

Qy    313 AGTAGCTGGGACTACAAGCTCGGGACACCAC 343
          |||||||||||||||||  | | ||||||
Db   1783 AGTAGCTGGGACTACAAGCATGCGCCACCAC 1813
```

RESULT 8
AAL06031
ID AAL06031 standard; DNA; 4045 BP.

XX
AC AAL06031;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen DNA SEQ ID NO: 8719.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200155320-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01339.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.

```
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-465570/50.
XX
PT   Isolated nucleic acid molecule encoding a reproductive system antigen
PT   is used in preventing, treating or ameliorating a medical condition -
XX
PS   Disclosure; SEQ ID NO 8719; 1297pp + Sequence Listing; English.
XX
CC   The present invention provides the protein and coding sequences of a
CC   number of human reproductive system related antigens. These can be used
CC   in the prevention and treatment of reproductive system disorders,
CC   including cancer. The present sequence is a genomic sequence encoding a
CC   protein of the invention.
XX
SQ   Sequence 4045 BP; 1044 A; 994 C; 1001 G; 1006 T; 0 other;

  Query Match             24.0%;  Score 119.8;  DB 22;  Length 4045;
  Best Local Similarity   73.0%;  Pred. No. 1.8e-22;
  Matches  154;  Conservative    0;  Mismatches   57;  Indels    0;  Gaps    0;

Qy     133 TTAGAGCTTGTCCTCTATTCTTGCTTTTCTTTCCAAAAAACACTACAATTTTTGTTTTGT 192
           ||  |  ||| | ||| |        ||| ||||        || |  | ||| | |
Db    1603 TTCTACATTGGCTTCTTTCTCGTTCCTTCCTTCCTTCCCTTCCTTCCTTCTTTCTCTCTC 1662

Qy     193 TTTGTTTTGTTTTGAGACAGGGTCTCGAGGTGTCACCCAGGCTGGAGTGCAGTGGCGCGA 252
           | | |||  |||||||||| |||||    |||| ||||||||||||||||||||| |||
Db    1663 TCTCTTTCTTTTTGAGACAGAGTCTCACTCTGTCGCCCAGGCTGGAGTGCAGTGGCACGA 1722

Qy     253 TTTCGACTCACCGCAACCTCCGCCTCCGCGCTTAAGCGATTCTCCTGCCTCAGCCTCCCA 312
           | ||  ||||| |||||||| ||||| | | | ||||||||||||||||||||||||||
Db    1723 TCTCAGCTCACTGCAACCTCTGCCTCTGAGTTCAAGCGATTCTCCTGCCTCAGCCTCCCA 1782

Qy     313 AGTAGCTGGGACTACAAGCTCGGGACACCAC 343
           |||||||||||||||||  | | ||||||
Db    1783 AGTAGCTGGGACTACAAGCATGCGCCACCAC 1813


RESULT 9
ABL98595
ID   ABL98595 standard; DNA; 4045 BP.
XX
AC   ABL98595;
XX
DT   21-JUN-2002  (first entry)
XX
DE   Human testicular antigen encoding DNA fragment SEQ ID NO: 3247.
XX
KW   Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW   reproductive system disorder; urinary system disorder; gene therapy;
KW   cardiovascular disorder; respiratory disorder; neurological disorder;
KW   gastrointestinal disease; infection; cytostatic; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200155317-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01329.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
```

PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483232/52.
XX
PT Nucleic acids encoding 973 human testicular antigen polypeptides,
PT useful for preventing, diagnosing and/or treating testicular cancer
XX
PS Disclosure; SEQ ID NO 3247; 766pp; English.
XX
CC The present invention provides the protein and coding sequences of 973
CC human testicular antigens, and fragments of their genomic sequences. The
CC sequences can be used in the treatment of cardiovascular, urinary system,
CC reproductive system, immune, respiratory, neurological and
CC gastrointestinal disorders, infections, and particularly cancer,
CC especially testicular cancers. The present sequence is a DNA encoding a
CC protein fragment of the invention.
XX
SQ Sequence 4045 BP; 1044 A; 994 C; 1001 G; 1006 T; 0 other;

Query Match 24.0%; Score 119.8; DB 23; Length 4045;
Best Local Similarity 73.0%; Pred. No. 1.8e-22;
Matches 154; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

```
Qy    133 TTAGAGCTTGTCCTCTATTCTTGCTTTTCTTTCCAAAAAACACTACAATTTTTGTTTTGT 192
          ||  |  ||| | ||| |        ||| ||||        || |  | ||| | |
Db   1603 TTCTACATTGGCTTCTTTCTCGTTCCTTCCTTCCTTCCCTTCCTTCCTTCTTTCTCTCTC 1662

Qy    193 TTTGTTTTGTTTTGAGACAGGGTCTCGAGGTGTCACCCAGGCTGGAGTGCAGTGGCGCGA 252
          | | |||  ||||||||||| |||||    |||| ||||||||||||||||||||| |||
Db   1663 TCTCTTTCTTTTTGAGACAGAGTCTCACTCTGTCGCCCAGGCTGGAGTGCAGTGGCACGA 1722

Qy    253 TTTCGACTCACCGCAACCTCCGCCTCCGCGCTTAAGCGATTCTCCTGCCTCAGCCTCCCA 312
          | ||  ||||| |||||||| ||||| | | | ||||||||||||||||||||||||||
Db   1723 TCTCAGCTCACTGCAACCTCTGCCTCTGAGTTCAAGCGATTCTCCTGCCTCAGCCTCCCA 1782

Qy    313 AGTAGCTGGGACTACAAGCTCGGGACACCAC 343
          |||||||||||||||||  | | ||||||
Db   1783 AGTAGCTGGGACTACAAGCATGCGCCACCAC 1813
```

RESULT 10
ABL98596
ID ABL98596 standard; DNA; 4045 BP.
XX
AC ABL98596;
XX
DT 21-JUN-2002 (first entry)
XX
DE Human testicular antigen encoding DNA fragment SEQ ID NO: 3248.
XX
KW Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW reproductive system disorder; urinary system disorder; gene therapy;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disease; infection; cytostatic; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200155317-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01329.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483232/52.
XX
PT Nucleic acids encoding 973 human testicular antigen polypeptides,
PT useful for preventing, diagnosing and/or treating testicular cancer
XX
PS Disclosure; SEQ ID NO 3248; 766pp; English.
XX
CC The present invention provides the protein and coding sequences of 973
CC human testicular antigens, and fragments of their genomic sequences. The
CC sequences can be used in the treatment of cardiovascular, urinary system,
CC reproductive system, immune, respiratory, neurological and
CC gastrointestinal disorders, infections, and particularly cancer,
CC especially testicular cancers. The present sequence is a DNA encoding a
CC protein fragment of the invention.
XX
SQ Sequence 4045 BP; 1044 A; 994 C; 1001 G; 1006 T; 0 other;

Query Match 24.0%; Score 119.8; DB 23; Length 4045;
Best Local Similarity 73.0%; Pred. No. 1.8e-22;
Matches 154; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Qy 133 TTAGAGCTTGTCCTCTATTCTTGCTTTTCTTTCCAAAAAACACTACAATTTTTGTTTTGT 192

```
Db   1603 TTCTACATTGGCTTCTTTCTCGTTCCTTCCTTCCTTCCCTTCCTTCCTTCTTTCTCTCTC 1662

Qy    193 TTTGTTTTGTTTTGAGACAGGGTCTCGAGGTGTCACCCAGGCTGGAGTGCAGTGGCGCGA 252
          | | |||  |||||||||| |||||    |||| |||||||||||||||||||||| |||
Db   1663 TCTCTTTCTTTTTGAGACAGAGTCTCACTCTGTCGCCCAGGCTGGAGTGCAGTGGCACGA 1722

Qy    253 TTTCGACTCACCGCAACCTCCGCCTCCGCGCTTAAGCGATTCTCCTGCCTCAGCCTCCCA 312
          | ||  ||||| |||||||| ||||| | | | |||||||||||||||||||||||||||
Db   1723 TCTCAGCTCACTGCAACCTCTGCCTCTGAGTTCAAGCGATTCTCCTGCCTCAGCCTCCCA 1782

Qy    313 AGTAGCTGGGACTACAAGCTCGGGACACCAC 343
          |||||||||||||||||  | | ||||||
Db   1783 AGTAGCTGGGACTACAAGCATGCGCCACCAC 1813


RESULT 11
AAH24652/c
ID   AAH24652 standard; DNA; 33780 BP.
XX
AC   AAH24652;
XX
DT   06-AUG-2001  (first entry)
XX
DE   Nucleotide sequence of disrupted in schizophrenia 1 (DIS1) intron 8.
XX
KW   Disrupted in schizophrenia 1 gene; DIS1 gene; schizophrenia;
KW   psychiatric disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200140301-A2.
XX
PD   07-JUN-2001.
XX
PF   28-NOV-2000; 2000WO-EP11915.
XX
PR   01-DEC-1999;   99EP-0309667.
XX
PA   (ALKU ) AKZO NOBEL NV.
PA   (MEDI-) MEDICAL RES COUNCIL.
PA   (UYED-) UNIV EDINBURGH.
XX
PI   Porteous D,  Millar K,  Blackwood D;
XX
DR   WPI; 2001-374796/39.
XX
PT   Novel isolated polynucleotide which surrounds a breakpoint on
PT   chromosome 1 involved in a balanced t(1;11) (q42.1;q14.3)
PT   translocation, and its encoded proteins, useful as medicament for
PT   treating psychiatric disorders  -
XX
PS   Disclosure; Page 39-50; 51pp; English.
XX
CC   The present sequence represents intron 8 of a human disrupted in
CC   schizophrenia 1 (DIS1) gene. DIS1 is disrupted by a (1:11)(q42.1:q14.3)
CC   translocation which segregates with schizophrenia. The DIS1 gene consists
CC   of 13 exons. Exon 11 contains an alternative splice site which removed 66
CC   nucleotides from the open reading frame. The final intron of DIS1 belongs
CC   to the rare AT-AC class of introns. DIS1 polypeptide fragments and
CC   antibodies are useful for in vitro diagnosis of a psychiatric disorder.
CC   DIS1 protein is useful in a screening assay for identifying new drugs for
CC   treating psychiatric disorders. DIS1 polypeptides and polynucleotides are
CC   useful for preventing and/or treating diseases related to schizophrenia.
XX
SQ   Sequence 33780 BP; 9140 A; 6838 C; 7205 G; 10408 T; 189 other;

  Query Match             24.0%;  Score 119.8;  DB 22;  Length 33780;
  Best Local Similarity   75.9%;  Pred. No. 3.6e-22;
  Matches  148;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy    160 TCTTTCCAAAAAACACTACAATTTTTGTTTTGTTTTGTTTTGTTTTGAGACAGGGTCTCG 219
          || |  ||| | ||  |    ||||| |||| |||| |||| ||||||||| | |||| |
Db  12701 TCCTAACAAGACACTGTGGCTTTTTTTTTTTTTTTTTTTTTTTTTGAGACGGAGTCTTG 12642

Qy    220 AGGTGTCACCCAGGCTGGAGTGCAGTGGCGCGATTTCGACTCACCGCAACCTCCGCCTCC 279
          ||||||||||||||||||| |||||||||| ||| |||||||||| ||||||||||
Db  12641 CTCTGTCACCCAGGCTGGAGTGTAGTGGCGCGATATCGGCTCACCGCAAGCTCCGCCTCC 12582

Qy    280 GCGCTTAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACTACAAGCTCGGGACA 339
          | | | || |||||||||||| ||||||| ||||| |||||||||| || |  | ||
Db  12581 CGGTTCATGCCATTCTCCTGCCTCCGCCTCCCGAGTAGATGGGACTACAGGCGCCCGCCA 12522

Qy    340 CCACGTAAAAATGAT 354
          |||  |  | ||
Db  12521 TGACGCCTAGCTAAT 12507


RESULT 12
AAQ85372/c
ID   AAQ85372 standard; cDNA; 1160 BP.
XX
AC   AAQ85372;
XX
DT   29-AUG-1995  (first entry)
XX
DE   RANTES cDNA sequence.
XX
KW   RANTES; T-cell; T-lymphocyte; heparanase; heparin; heparan sulfate;
KW   arthritis; restenosis; cancer; wound healing; ss.
XX
OS   Homo sapiens.
XX
PN   WO9504158-A.
XX
PD   09-FEB-1995.
XX
PF   26-JUL-1994;   94WO-US08207.
XX
PR   29-JUL-1993;   93US-0099866.
PR   13-OCT-1993;   93US-0136117.
XX
PA   (UPJO ) UPJOHN CO.
XX
PI   Hoogwerf AJ,  Ledbetter SR;
XX
DR   WPI; 1995-082239/11.
DR   P-PSDB; AAR70802.
XX
PT   Screening for cpds. with anti-heparanase activity - by detecting
PT   inhibition of heparin or heparan sulphate degradation,
PT   potentially useful for treating arthritis, restenosis, cancer.
XX
PS   Disclosure; Page 51-52; 60pp; English.
XX
CC   Purified heparanases, prepared under reducing conditions and
CC   activated with transglutaminase, are given in AAR70786-805.  Most
CC   are prepared by reverse transcription of mRNA from activated human
CC   leukocytes, then cloning of the cDNA into pVL1392 baculovirus
CC   vector, and expression in Sf9 cells in the presence of reduced
CC   glutathione and dithiothreitol.
XX
SQ   Sequence 1160 BP; 298 A; 332 C; 295 G; 235 T; 0 other;

  Query Match             23.9%;  Score 119.6;  DB 16;  Length 1160;
  Best Local Similarity   82.5%;  Pred. No. 1.3e-22;
  Matches  137;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy    179 AATTTTTGTTTTGTTTTGTTTTGTTTTGAGACAGGGTCTCGAGGTGTCACCCAGGCTGGA 238
          | ||||| |||| |||| |||| ||||||||| | ||||||   |||| ||||||||||
Db    767 ATTTTTTTTTTTTTTTTTTTTTTTTTGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGA 708

Qy    239 GTGCAGTGGCGCGATTTCGACTCACCGCAACCTCCGCCTCCGCGCTTAAGCGATTCTCCT 298
          ||||||||||||||| ||| ||||| |||| ||||||  ||| | | | || ||||||||
Db    707 GTGCAGTGGCGCGATCTCGGCTCACTGCAAGCTCCGCTCCCGGGTTCACGCCATTCTCCT 648

Qy    299 GCCTCAGCCTCCCAAGTAGCTGGGACTACAAGCTCGGGACACCACG 344
```

```
          |||||||||||||| |||||||||||||||| || |  |  ||||||
Db    647 GCCTCAGCCTCCCGAGTAGCTGGGACTACAGGCGCCCGCTACCACG 602


RESULT 13
AAV36275/c
ID   AAV36275 standard; DNA; 1160 BP.
XX
AC   AAV36275;
XX
DT   01-SEP-1998  (first entry)
XX
DE   cDNA containing the sequence coding for RANTES.
XX
KW   Chemokine; RANTES; antagonist; antiviral; HIV; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             96..302
FT                   /*tag= a
FT                   /product= RANTES
XX
PN   WO9744462-A1.
XX
PD   27-NOV-1997.
XX
PF   22-MAY-1997;   97WO-FR00900.
XX
PR   22-MAY-1996;   96FR-0006368.
XX
PA   (INSP ) INST PASTEUR.
XX
PI   Bachelerie F,  Baggiolini M,  Virelizier JL,  Arenzana-Seisdedos F;
PI   Clark-Lewis I;
XX
DR   WPI; 1998-018516/02.
DR   P-PSDB; AAW29538.
XX
PT   Using chemokine antagonists for treating or preventing viral
PT   infection - particularly by human immunodeficiency virus, also new
PT   polypeptide antagonists derived from RANTES and related nucleic acid
XX
PS   Disclosure; Fig 1A; 41pp; French.
XX
CC   The invention relates to the use of an antagonist of one or more
CC   chemokines for treatment or prevention of viral infection. In particular
CC   the viral infection is HIV and the antagonist is a peptide composed of
CC   amino acids 8-68 or 9-68 of RANTES, or a substitution or deletion
CC   analogue of this peptide. The peptide may be produced by recombinant
CC   DNA techniques. The antagonist is especially used to treat HIV in
CC   children born to seropositive mothers or in cases of accidental
CC   contamination. It reduces replication and/or spread of the
CC   virus and/or corrects immunodeficiency caused by infection. The present
CC   sequence is the DNA coding for RANTES.
XX
SQ   Sequence 1160 BP; 298 A; 332 C; 295 G; 235 T; 0 other;

  Query Match             23.9%;  Score 119.6;  DB 19;  Length 1160;
  Best Local Similarity   82.5%;  Pred. No. 1.3e-22;
  Matches  137;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy    179 AATTTTTGTTTTGTTTTGTTTTGTTTTGAGACAGGGTCTCGAGGTGTCACCCAGGCTGGA 238
          | ||||| |||| |||| |||| ||||||||| | ||||||   |||| ||||||||||||
Db    767 ATTTTTTTTTTTTTTTTTTTTTTTTTTTGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGA 708

Qy    239 GTGCAGTGGCGCGATTTCGACTCACCGCAACCTCCGCCTCCGCGCTTAAGCGATTCTCCT 298
          ||||||||||||||| ||| ||||| |||| ||||||  ||| | | | || ||||||||
Db    707 GTGCAGTGGCGCGATCTCGGCTCACTGCAAGCTCCGCTCCCGGGTTCACGCCATTCTCCT 648

Qy    299 GCCTCAGCCTCCCAAGTAGCTGGGACTACAAGCTCGGGACACCACG 344
          |||||||||||||| |||||||||||||||| || |  |  ||||||
Db    647 GCCTCAGCCTCCCGAGTAGCTGGGACTACAGGCGCCCGCTACCACG 602


RESULT 14
AAF21050/c
ID   AAF21050 standard; DNA; 1160 BP.
XX
AC   AAF21050;
XX
DT   14-MAR-2001  (first entry)
XX
DE   Human low adenosine antisense oligonucleotide related sequence #2617.
XX
KW   Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
KW   human; airway disorder; bronchoconstriction; lung inflammation;
KW   surfactant depletion; respiratory; bronchodilator; antiinflammatory;
KW   immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
KW   respiratory obstruction; pulmonary obstruction; impeded respiration;
KW   surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
KW   respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
KW   pulmonary hypertension; emphysema; pulmonary transplantation rejection;
KW   chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
KW   cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200062736-A2.
XX
PD   26-OCT-2000.
XX
PF   24-MAR-2000; 2000WO-US08020.
XX
PR   06-APR-1999;   99US-0127958.
XX
PA   (UYEC-) UNIV EAST CAROLINA.
PA   (NYCE/) NYCE J W.
XX
PI   Nyce JW;
XX
DR   WPI; 2000-679539/66.
XX
PT   Low adenosine (A) content antisense oligonucleotides which do not
PT   trigger adenosine receptors during metabolism, useful e.g. for treating
PT   cancers and respiratory obstructions -
XX
PS   Disclosure; Page 856-857; 1592pp; English.
XX
CC   The present invention describes low adenosine (A) content antisense
CC   oligonucleotides and compositions (I) comprising them. In the antisense
CC   oligonucleotides the A is replaced by a 'Universal' or alternative base.
CC   (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
CC   immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
CC   The antisense oligonucleotides and (I) can be used to down-regulate the
CC   expression and or activity of target polypeptides associated with
CC   lung/respiratory disorders and malignancies, such as stimulating and
CC   activating peptide factors and transmitters, transcription factors,
CC   immunoglobulins and antibodies, antibody receptors, cytokines and
CC   chemokines, endogenously produced specific and non-specific enzymes,
CC   binding proteins, adhesion molecules and their receptors, cytokine and
CC   chemokine receptors, adenosine receptors, bradykinin receptors, central
CC   nervous system (CNS) and peripheral nervous and non-nervous system
CC   receptors, CNS and peripheral nervous and non-nervous system peptide
CC   transmitters, defensins, growth factors, vasoactive peptides and
CC   receptors, binding proteins and malignancy associated proteins. The
CC   antisense oligonucleotides may be used in this way to treat disorders
CC   including respiratory obstruction (especially pulmonary obstruction
CC   and/or bronchoconstriction) and/or lung inflammation, allergy(ies)
CC   and/or surfactant hypoproduction which are associated with a disease or
CC   condition selected from pulmonary vasoconstriction, inflammation,
CC   allergies, asthma, impeded respiration, respiratory distress syndrome
CC   (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC   hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
CC   pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC   and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
```

```
CC   fragments and antisense oligonucleotides used in the exemplification of
CC   the present invention.
XX
SQ   Sequence 1160 BP; 298 A; 332 C; 295 G; 235 T; 0 other;

  Query Match             23.9%;  Score 119.6;  DB 21;  Length 1160;
  Best Local Similarity   82.5%;  Pred. No. 1.3e-22;
  Matches  137;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy     179 AATTTTTGTTTTGTTTTGTTTTGTTTTGAGACAGGGTCTCGAGGTGTCACCCAGGCTGGA 238
           | ||||| |||| |||| |||| ||||||||| | ||||||   |||| ||||||||||
Db     767 ATTTTTTTTTTTTTTTTTTTTTTTTTTGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGA 708

Qy     239 GTGCAGTGGCGCGATTTCGACTCACCGCAACCTCCGCCTCCGCGCTTAAGCGATTCTCCT 298
           ||||||||||||||| ||| ||||| |||| |||||| ||| | | | || ||||||||
Db     707 GTGCAGTGGCGCGATCTCGGCTCACTGCAAGCTCCGCTCCCGGGTTCACGCCATTCTCCT 648

Qy     299 GCCTCAGCCTCCCAAGTAGCTGGGACTACAAGCTCGGGACACCACG 344
           |||||||||||| |||||||||||||||| || |  |  ||||||
Db     647 GCCTCAGCCTCCCGAGTAGCTGGGACTACAGGCGCCCGCTACCACG 602


RESULT 15
AAA74884/c
ID   AAA74884 standard; DNA; 1160 BP.
XX
AC   AAA74884;
XX
DT   17-JAN-2001  (first entry)
XX
DE   Human chemokine coding sequence SEQ ID NO: 78.
XX
KW   Macrophage recruitment; chemokine derivative; MCP-1; osteoporosis;
KW   monocyte chemoattractant protein-1; inflammation; atherosclerosis; HIV;
KW   AIDS; stroke; psoriasis; autoimmune disease; hypertension; endotoxaemia;
KW   basophil-mediated disease; myocardial infarction; acute ischaemia;
KW   rheumatoid arthritis; contraception; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             27..302
FT                   /*tag=  a
FT                   /product= "human chemokine"
XX
PN   WO200042071-A2.
XX
PD   20-JUL-2000.
XX
PF   12-JAN-2000; 2000WO-US00821.
XX
PR   12-JAN-1999;   99US-0229071.
PR   17-MAR-1999;   99US-0271192.
PR   01-DEC-1999;   99US-0452406.
XX
PA   (NEOR-) NEORX CORP.
XX
PI   Grainger DJ,  Tatalick LM;
XX
DR   WPI; 2000-499101/44.
DR   P-PSDB; AAB15790.
XX
PT   New peptide 3, amide and heterocyclic compounds and saccharide
PT   conjugates used for inhibiting chemokine induced activity and for
PT   treating e.g. stroke, vascular diseases, autoimmune diseases and tumour
PT   growth  -
XX
PS   Disclosure; Page 375-376; 387pp; English.
XX
CC   The present invention concerns the identification of a number of
CC   chemokines which can be used to produce derivatives, agonists and
CC   antagonists which are then useful in disease treatment. The chemokines
CC   include sequences AAB15785-B15794, AAB15803-B15813 and AAB15831-B15848.
CC   These chemokine derivatives can be used to treat diseases such as
CC   autoimmune diseases, atherosclerosis, osteoporosis, HIV infection and
CC   AIDS, psoriasis, inflammatory diseases, hypertension, basophil-mediated
CC   diseases, endotoxaemia, myocardial infarction, acute ischaemia and
CC   rheumatoid arthritis, and can be used to prevent strokes and as
CC   contraceptives. The chemokine coding sequences AAA74868-A74888 can be
CC   used in gene therapy for the same diseases, as well as in the production
CC   of animal models.
XX
SQ   Sequence 1160 BP; 298 A; 332 C; 295 G; 235 T; 0 other;

  Query Match             23.9%;  Score 119.6;  DB 21;  Length 1160;
  Best Local Similarity   82.5%;  Pred. No. 1.3e-22;
  Matches  137;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy     179 AATTTTTGTTTTGTTTTGTTTTGTTTTGAGACAGGGTCTCGAGGTGTCACCCAGGCTGGA 238
           | ||||| |||| |||| |||| ||||||||| | ||||||   |||| ||||||||||
Db     767 ATTTTTTTTTTTTTTTTTTTTTTTTTTGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGA 708

Qy     239 GTGCAGTGGCGCGATTTCGACTCACCGCAACCTCCGCCTCCGCGCTTAAGCGATTCTCCT 298
           ||||||||||||||| ||| ||||| |||| |||||| ||| | | | || ||||||||
Db     707 GTGCAGTGGCGCGATCTCGGCTCACTGCAAGCTCCGCTCCCGGGTTCACGCCATTCTCCT 648

Qy     299 GCCTCAGCCTCCCAAGTAGCTGGGACTACAAGCTCGGGACACCACG 344
           |||||||||||| |||||||||||||||| || |  |  ||||||
Db     647 GCCTCAGCCTCCCGAGTAGCTGGGACTACAGGCGCCCGCTACCACG 602


Search completed: May 16, 2003, 03:39:40
Job time : 1159.37 secs
```

GenCore version 5.1.5

OM nucleic - nucleic search, using sw model

Run on: May 16, 2003, 01:11:33 ; Search time 111.593 Seconds
(without alignments)
10110.387 Million cell updates/sec

Title: US-09-434-382-28_COPY_26164_26664
Perfect score: 501
Sequence: 1 ggtatggagctgtgccgagg.....ttcgccaagtctttttgaca 501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	501	100.0	26664	21	AAA60207	Human prostate can
	2	501	100.0	26664	24	AAS98942	Human prostate can
c	3	501	100.0	37959	22	AAK81763	Human immune/haema
	4	303	60.5	655	24	AAS98941	Human prostate can
	5	303	60.5	2958	21	AAA58453	Human prostate can
	6	303	60.5	2958	24	AAS98917	Human prostate can
	7	299.8	59.8	2546	21	AAC76445	Human ORFX ORF2000
	8	296.6	59.2	2908	24	AAS99132	Chimpanzee ELAC2 c
	9	296.6	59.2	2992	24	ABN59829	Novel human coding
	10	294.6	58.8	2976	22	AAH14250	Human cDNA sequenc
c	11	293.6	58.6	481	22	AAH10926	Human cDNA clone (
	12	254.6	50.8	2892	24	AAS99133	Gorilla ELAC2 cDNA
	13	204	40.7	1402	23	AAS72207	DNA encoding novel
	14	158	31.5	386	22	AAF64286	Novel human polynu
	15	155	30.9	1450	23	AAS72208	DNA encoding novel
c	16	134.4	26.8	137	20	AAH85145	Human single nucle
c	17	95.4	19.0	97	24	ABK17279	Coupled ligation a
	18	95.4	19.0	97	24	ABK17312	Coupled ligation a
c	19	93.8	18.7	97	24	ABK17278	Coupled ligation a
	20	93.8	18.7	97	24	ABK17299	Coupled ligation a
	21	70	14.0	72	16	AAT25953	Human gene signatu
c	22	65.4	13.1	68	19	AAX11554	Human biallelic po
	23	60	12.0	60	24	ABN32391	Human spliced tran
c	24	34.6	6.9	1455	21	AAC36566	Arabidopsis thalia
	25	33.8	6.7	1596	22	AAK65307	Human immune/haema
c	26	33.6	6.7	480	22	ABA42670	Human breast cell
c	27	33.6	6.7	480	22	AAI01349	Probe #1340 used t
c	28	33.6	6.7	480	24	ABS01397	Human genome-deriv
	29	33.4	6.7	5959	24	ABQ67080	Human angiogenesis
	30	33.4	6.7	5959	24	ABL33281	Human immune syste
	31	33.2	6.6	6042	17	AAT09018	Arabidopsis thalia
	32	33.2	6.6	6172	19	AAV57454	Arabidopsis ethyle
	33	33.2	6.6	6172	22	AAD03789	Arabidopsis thalia
c	34	32.6	6.5	1425	23	AAS80594	DNA encoding novel
	35	32.2	6.4	626	24	ABN87200	Lolium perenne LpC
c	36	32.2	6.4	633	24	ABQ69328	Listeria innocua D
c	37	32.2	6.4	654	24	ABQ67637	Listeria innocua D
	38	32.2	6.4	777	24	ABN87198	Lolium perenne LpC
	39	32.2	6.4	513445	22	AAI61373	Soybean 318013 reg
	40	32.2	6.4	684707	24	ABQ67196	Listeria innocua c
c	41	32.2	6.4	3011208	24	ABQ69245	Listeria innocua D
	42	32	6.4	500	22	AAH12149	Human cDNA clone (
c	43	32	6.4	1089	24	ABK75029	Bacillus lichenifo
c	44	32	6.4	1984	22	AAH16379	Human cDNA sequenc
	45	32	6.4	3779	24	ABN95647	Gene #2145 used to

ALIGNMENTS

RESULT 1
AAA60207
ID AAA60207 standard; DNA; 26664 BP.
XX
AC AAA60207;
XX
DT 07-DEC-2000 (first entry)
XX
DE Human prostate cancer predisposing gene HPC2 genomic sequence.
XX
KW Human; prostate cancer predisposing gene; HPC2; chromosome 17p;
KW gene therapy; peptide therapy; drug design; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 910..26039
FT /*tag= a
FT /product= "HPC2"
FT /note= "this sequence contains introns"
FT /transl_except= (pos:23892..23895,aa:Glu)
FT exon 910..1154
FT /*tag= b
FT /number= 1
FT exon 1736..1786
FT /*tag= c
FT /number= 2
FT exon 1925..1995
FT /*tag= d
FT /number= 3
FT exon 3025..3089
FT /*tag= e

```
FT                   /number= 4
FT   exon            4361..4418
FT                   /*tag=  f
FT                   /number= 5
FT   exon            5582..5650
FT                   /*tag=  g
FT                   /number= 6
FT   exon            7075..7194
FT                   /*tag=  h
FT                   /number= 7
FT   exon            8186..8244
FT                   /*tag=  i
FT                   /number= 8
FT   exon            12878..12936
FT                   /*tag=  j
FT                   /number= 9
FT   exon            13032..13104
FT                   /*tag=  k
FT                   /number= 10
FT   exon            13756..13868
FT                   /*tag=  l
FT                   /number= 11
FT   exon            15283..15378
FT                   /*tag=  m
FT                   /number= 12
FT   exon            16278..16416
FT                   /*tag=  n
FT                   /number= 13
FT   exon            16498..16583
FT                   /*tag=  o
FT                   /number= 14
FT   exon            18583..18701
FT                   /*tag=  p
FT                   /number= 15
FT   exon            20349..20445
FT                   /*tag=  q
FT                   /number= 16
FT   exon            22172..22310
FT                   /*tag=  r
FT                   /number= 17
FT   exon            22879..22917
FT                   /*tag=  s
FT                   /number= 18
FT   exon            23045..23154
FT                   /*tag=  t
FT                   /number= 19
FT   exon            23795..23895
FT                   /*tag=  u
FT                   /number= 20
FT   exon            23973..24093
FT                   /*tag=  v
FT                   /number= 21
FT   exon            24354..24432
FT                   /*tag=  w
FT                   /number= 22
FT   exon            25026..25170
FT                   /*tag=  x
FT                   /number= 23
FT   exon            25812..26036
FT                   /*tag=  y
FT                   /number= 24
FT   polyA_signal    26447..26452
FT                   /*tag=  z
XX
PN   WO200027864-A1.
XX
PD   18-MAY-2000.
XX
PF   05-NOV-1999;   99WO-US26055.
XX
PR   06-NOV-1998;   98US-0107468.
XX
PA   (MYRI-) MYRIAD GENETICS INC.
XX
PI   Tavtigian SV,  Teng DHF,  Simard J,  Rommens JM;
XX
DR   WPI; 2000-376481/32.
DR   P-PSDB; AAB07228.
XX
PT   Human prostate cancer (HPC)2 nucleic acids, polypeptides, and
PT   antibodies, useful for treatment and diagnosis of prostate cancer -
XX
PS   Claim 3; Page 108-122; 157pp; English.
XX
CC   The present sequence is the genomic sequence of the human prostate
CC   cancer predisposing gene HPC2, which is found on chromosome 17p. Some
CC   alleles of this gene cause a predisposition to cancer, particularly
CC   prostate cancer. This gene and its protein can be used in peptide and
CC   gene therapy for cancer patients, as well as being useful as diagnostic
CC   tools (both for cancer sufferers and those with a predisposition to the
CC   disease) and in the production of cancer drugs. This sequence was
CC   isolated by cloning and sequencing the region of the genome which
CC   appeared to cause a predisposition to prostate cancer.
XX
SQ   Sequence 26664 BP; 6173 A; 6300 C; 6519 G; 7661 T; 11 other;

  Query Match             100.0%;  Score 501;  DB 21;  Length 26664;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-157;
  Matches  501;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26164 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTT 26223

Qy      61 AGGACTGGTGCCTGGCACAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26224 AGGACTGGTGCCTGGCACAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 26283

Qy     121 ACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCTTGGAAACAGACGGCGGCACCTTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26284 ACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCTTGGAAACAGACGGCGGCACCTTTC 26343

Qy     181 CTCTAATCCAGCAAAGTGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26344 CTCTAATCCAGCAAAGTGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTG 26403

Qy     241 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26404 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 26463

Qy     301 CAATTGTGAGTTCTTTTGCTTCCTCCTGCTGCTGCTACAGAGCAGGGTCTGCTGTGCACC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26464 CAATTGTGAGTTCTTTTGCTTCCTCCTGCTGCTGCTACAGAGCAGGGTCTGCTGTGCACC 26523

Qy     361 ACCTTGGAGAAGGCTCTCTGTGCTGTAGTGTGGCAGCTGCCTGGTACCCGGGTGGCTTGG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26524 ACCTTGGAGAAGGCTCTCTGTGCTGTAGTGTGGCAGCTGCCTGGTACCCGGGTGGCTTGG 26583

Qy     421 AAGAAGTCAGCTCCCGTCGTAGTGAGCACCTCTGGAACCTGTCCTCAGAGAGCCACCCTT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26584 AAGAAGTCAGCTCCCGTCGTAGTGAGCACCTCTGGAACCTGTCCTCAGAGAGCCACCCTT 26643

Qy     481 ATTCGCCAAGTCTTTTTGACA 501
          |||||||||||||||||||||
Db   26644 ATTCGCCAAGTCTTTTTGACA 26664


RESULT 2
AAS98942
ID   AAS98942 standard; DNA; 26664 BP.
XX
AC   AAS98942;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Human prostate cancer predisposing gene (HPC2) genomic DNA.
```

XX
KW Human; mouse; HPC2; prostate cancer; neoplastic growth; cytostatic; ds;
KW gene therapy; prostate cancer predisposing gene.
XX
OS Homo sapiens.
XX
PN WO200185911-A2.
XX
PD 15-NOV-2001.
XX
PF 07-MAY-2001; 2001WO-US14602.
XX
PR 05-MAY-2000; 2000US-0564805.
XX
PA (MYRI-) MYRIAD GENETICS INC.
PA (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI Tavtigian SV, Teng DHF, Simard J, Rommens JM;
XX
DR WPI; 2002-066599/09.
XX
PT Novel nucleic acid sequence encoding HPC2 polypeptide, which is marker
PT for prostate cancer, is useful in gene therapy techniques to restore
PT HPC2 normal levels by which neoplastic growth is suppressed in
PT recipient cell
XX
PS Claim 3; Page 143-157; 239pp; English.
XX
CC The invention relates to a human prostate cancer predisposing gene coding
CC for an HPC2 polypeptide. The DNA and protein sequences are useful as
CC diagnostic reagents for identifying a mutant HPC2 nucleotide sequence in
CC a suspected mutant HPC2 allele by comparing the sequence of the suspected
CC mutant HPC2 allele with a wild-type HPC2 sequence. The sequences are also
CC useful for detecting an alteration in HPC2, where the alteration is
CC associated with cancer in a human. The method involves analysing an HPC2
CC gene or an HPC2 gene expression product from a tissue of the human. The
CC HPC2 gene is useful as a marker for prostate cancer and can be used in
CC gene therapy techniques to suppress neoplastic growth of recipient cells
CC which carry the mutant HPC2 allele. The sequences represent DNA encoding
CC human and mouse HPC2 and fragments of HPC2.
XX
SQ Sequence 26664 BP; 6173 A; 6300 C; 6519 G; 7661 T; 11 other;

Query Match 100.0%; Score 501; DB 24; Length 26664;
Best Local Similarity 100.0%; Pred. No. 6.5e-157;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26164 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTT 26223

Qy      61 AGGACTGGTGCCTGGCACAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26224 AGGACTGGTGCCTGGCACAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 26283

Qy     121 ACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCTTGGAAACAGACGGCGGCACCTTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26284 ACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCTTGGAAACAGACGGCGGCACCTTTC 26343

Qy     181 CTCTAATCCAGCAAAGTGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26344 CTCTAATCCAGCAAAGTGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTG 26403

Qy     241 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26404 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 26463

Qy     301 CAATTGTGAGTTCTTTTGCTTCCTCCTGCTGCTGCTACAGAGCAGGGTCTGCTGTGCACC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26464 CAATTGTGAGTTCTTTTGCTTCCTCCTGCTGCTGCTACAGAGCAGGGTCTGCTGTGCACC 26523

Qy     361 ACCTTGGAGAAGGCTCTCTGTGCTGTAGTGTGGCAGCTGCCTGGTACCCGGGTGGCTTGG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26524 ACCTTGGAGAAGGCTCTCTGTGCTGTAGTGTGGCAGCTGCCTGGTACCCGGGTGGCTTGG 26583

Qy     421 AAGAAGTCAGCTCCCGTCGTAGTGAGCACCTCTGGAACCTGTCCTCAGAGAGCCACCCTT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26584 AAGAAGTCAGCTCCCGTCGTAGTGAGCACCTCTGGAACCTGTCCTCAGAGAGCCACCCTT 26643

Qy     481 ATTCGCCAAGTCTTTTTGACA 501
          |||||||||||||||||||||
Db   26644 ATTCGCCAAGTCTTTTTGACA 26664
```

RESULT 3
AAK81763/c
ID AAK81763 standard; DNA; 37959 BP.
XX
AC AAK81763;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:36575.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.

PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS Disclosure; SEQ ID NO 36575; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 37959 BP; 10440 A; 9397 C; 9111 G; 9011 T; 0 other;

Query Match 100.0%; Score 501; DB 22; Length 37959;
Best Local Similarity 100.0%; Pred. No. 7.9e-157;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12631 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTT 12572

Qy     61 AGGACTGGTGCCTGGCACAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12571 AGGACTGGTGCCTGGCACAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 12512

Qy    121 ACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCTTGGAAACAGACGGCGGCACCTTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12511 ACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCTTGGAAACAGACGGCGGCACCTTTC. 12452

Qy    181 CTCTAATCCAGCAAAGTGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  12451 CTCTAATCCAGCAAAGTGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTG 12392

Qy    241 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12391 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 12332

Qy    301 CAATTGTGAGTTCTTTTGCTTCCTCCTGCTGCTGCTACAGAGCAGGGTCTGCTGTGCACC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12331 CAATTGTGAGTTCTTTTGCTTCCTCCTGCTGCTGCTACAGAGCAGGGTCTGCTGTGCACC 12272

Qy    361 ACCTTGGAGAAGGCTCTCTGTGCTGTAGTGTGGCAGCTGCCTGGTACCCGGGTGGCTTGG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12271 ACCTTGGAGAAGGCTCTCTGTGCTGTAGTGTGGCAGCTGCCTGGTACCCGGGTGGCTTGG 12212

Qy    421 AAGAAGTCAGCTCCCGTCGTAGTGAGCACCTCTGGAACCTGTCCTCAGAGAGCCACCCTT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12211 AAGAAGTCAGCTCCCGTCGTAGTGAGCACCTCTGGAACCTGTCCTCAGAGAGCCACCCTT 12152

Qy    481 ATTCGCCAAGTCTTTTTGACA 501
          |||||||||||||||||||||
Db  12151 ATTCGCCAAGTCTTTTTGACA 12131
```

RESULT 4
AAS98941
ID AAS98941 standard; DNA; 655 BP.
XX
AC AAS98941;
XX
DT 12-MAR-2002 (first entry)
XX
DE Human prostate cancer predisposing gene (HPC2) DNA partial exon #24.
XX
KW Human; mouse; HPC2; prostate cancer; neoplastic growth; cytostatic; ds;
KW gene therapy; prostate cancer predisposing gene.
XX
OS Homo sapiens.
XX
PN WO200185911-A2.
XX
PD 15-NOV-2001.
XX
PF 07-MAY-2001; 2001WO-US14602.
XX
PR 05-MAY-2000; 2000US-0564805.
XX
PA (MYRI-) MYRIAD GENETICS INC.
PA (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI Tavtigian SV, Teng DHF, Simard J, Rommens JM;
XX
DR WPI; 2002-066599/09.
XX
PT Novel nucleic acid sequence encoding HPC2 polypeptide, which is marker
PT for prostate cancer, is useful in gene therapy techniques to restore
PT HPC2 normal levels by which neoplastic growth is suppressed in
PT recipient cell -
XX
PS Claim 9; Page 142; 239pp; English.
XX
CC The invention relates to a human prostate cancer predisposing gene coding
CC for an HPC2 polypeptide. The DNA and protein sequences are useful as
CC diagnostic reagents for identifying a mutant HPC2 nucleotide sequence in
CC a suspected mutant HPC2 allele by comparing the sequence of the suspected
CC mutant HPC2 allele with a wild-type HPC2 sequence. The sequences are also
CC useful for detecting an alteration in HPC2, where the alteration is
CC associated with cancer in a human. The method involves analysing an HPC2
CC gene or an HPC2 gene expression product from a tissue of the human. The
CC HPC2 gene is useful as a marker for prostate cancer and can be used in
CC gene therapy techniques to suppress neoplastic growth of recipient cells
CC which carry the mutant HPC2 allele. The sequences represent DNA encoding
CC human and mouse HPC2 and fragments of HPC2.
XX

SQ Sequence 655 BP; 165 A; 169 C; 199 G; 122 T; 0 other;

  Query Match   60.5%; Score 303; DB 24; Length 655;
  Best Local Similarity   100.0%; Pred. No. 3.2e-91;
  Matches 303; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy      1 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    353 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTT 412

Qy     61 AGGACTGGTGCCTGGCACAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    413 AGGACTGGTGCCTGGCACAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 472

Qy    121 ACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCTTGGAAACAGACGGCGGCACCTTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    473 ACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCTTGGAAACAGACGGCGGCACCTTTC 532

Qy    181 CTCTAATCCAGCAAAGTGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    533 CTCTAATCCAGCAAAGTGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTG 592

Qy    241 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    593 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 652

Qy    301 CAA 303
          |||
Db    653 CAA 655
```

RESULT 5
AAA58453
ID AAA58453 standard; cDNA; 2958 BP.
XX
AC AAA58453;
XX
DT 07-DEC-2000 (first entry)
XX
DE Human prostate cancer predisposing gene HPC2 coding sequence.
XX
KW Human; prostate cancer predisposing gene; HPC2; chromosome 17p;
KW gene therapy; peptide therapy; drug design; ss.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT CDS   51..2531
FT   /*tag= a
FT   /product= "HPC2"
XX
PN WO200027864-A1.
XX
PD 18-MAY-2000.
XX
PF 05-NOV-1999; 99WO-US26055.
XX
PR 06-NOV-1998; 98US-0107468.
XX
PA (MYRI-) MYRIAD GENETICS INC.
XX
PI Tavtigian SV, Teng DHF, Simard J, Rommens JM;
XX
DR WPI; 2000-376481/32.
DR P-PSDB; AAB07228.
XX
PT Human prostate cancer (HPC)2 nucleic acids, polypeptides, and
PT antibodies, useful for treatment and diagnosis of prostate cancer -
XX
PS Claim 3; Page 98-100; 157pp; English.
XX
CC The present sequence is the coding sequence of the human prostate
CC cancer predisposing gene HPC2, which is found on chromosome 17p. Some

```
CC   alleles of this gene cause a predisposition to cancer, particularly
CC   prostate cancer. This gene and its protein can be used in peptide and
CC   gene therapy for cancer patients, as well as being useful as diagnostic
CC   tools (both for cancer sufferers and those with a predisposition to the
CC   disease) and in the production of cancer drugs. This sequence was
CC   isolated by cloning and sequencing the region of the genome which
CC   appeared to cause a predisposition to prostate cancer.
XX
SQ   Sequence 2958 BP; 707 A; 805 C; 848 G; 598 T; 0 other;

  Query Match             60.5%;  Score 303;  DB 21;  Length 2958;
  Best Local Similarity   100.0%;  Pred. No. 7.4e-91;
  Matches  303;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2656 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTT 2715

Qy     61 AGGACTGGTGCCTGGCACAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2716 AGGACTGGTGCCTGGCACAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 2775

Qy    121 ACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCTTGGAAACAGACGGCGGCACCTTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2776 ACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCTTGGAAACAGACGGCGGCACCTTTC 2835

Qy    181 CTCTAATCCAGCAAAGTGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2836 CTCTAATCCAGCAAAGTGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTG 2895

Qy    241 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2896 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 2955

Qy    301 CAA 303
          |||
Db   2956 CAA 2958


RESULT 6
AAS98917
ID   AAS98917 standard; cDNA; 2958 BP.
XX
AC   AAS98917;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Human prostate cancer predisposing gene (HPC2) extended cDNA.
XX
KW   Human; mouse; HPC2; prostate cancer; neoplastic growth; cytostatic; ss;
KW   gene therapy; prostate cancer predisposing gene; chimpanzee; gorilla;
KW   sequencing primer; PCR primer.
XX
OS   Homo sapiens.
XX
PN   WO200185911-A2.
XX
PD   15-NOV-2001.
XX
PF   07-MAY-2001; 2001WO-US14602.
XX
PR   05-MAY-2000; 2000US-0564805.
XX
PA   (MYRI-) MYRIAD GENETICS INC.
PA   (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI   Tavtigian SV,  Teng DHF,  Simard J,  Rommens JM;
XX
DR   WPI; 2002-066599/09.
DR   P-PSDB; AAU73586.
XX
PT   Novel nucleic acid sequence encoding HPC2 polypeptide, which is marker
PT   for prostate cancer, is useful in gene therapy techniques to restore
PT   HPC2 normal levels by which neoplastic growth is suppressed in
PT   recipient cell -
XX
PS   Claim 3; Page 134-136; 239pp; English.
XX
CC   The invention relates to a human prostate cancer predisposing gene coding
CC   for an HPC2 polypeptide. The DNA and protein sequences are useful as
CC   diagnostic reagents for identifying a mutant HPC2 nucleotide sequence in
CC   a suspected mutant HPC2 allele by comparing the sequence of the suspected
CC   mutant HPC2 allele with a wild-type HPC2 sequence. The sequences are also
CC   useful for detecting an alteration in HPC2, where the alteration is
CC   associated with cancer in a human. The method involves analysing an HPC2
CC   gene or an HPC2 gene expression product from a tissue of the human. The
CC   HPC2 gene is useful as a marker for prostate cancer and can be used in
CC   gene therapy techniques to suppress neoplastic growth of recipient cells
CC   which carry the mutant HPC2 allele. The sequences represent primers used
CC   in the methods of the invention, cDNA encoding human and mouse HPC2 and
CC   cDNA encoding HPC2 paralogues and orthologues.
XX
SQ   Sequence 2958 BP; 707 A; 805 C; 848 G; 598 T; 0 other;

  Query Match             60.5%;  Score 303;  DB 24;  Length 2958;
  Best Local Similarity   100.0%;  Pred. No. 7.4e-91;
  Matches  303;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2656 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTT 2715

Qy     61 AGGACTGGTGCCTGGCACAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2716 AGGACTGGTGCCTGGCACAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 2775

Qy    121 ACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCTTGGAAACAGACGGCGGCACCTTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2776 ACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCTTGGAAACAGACGGCGGCACCTTTC 2835

Qy    181 CTCTAATCCAGCAAAGTGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2836 CTCTAATCCAGCAAAGTGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTG 2895

Qy    241 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2896 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 2955

Qy    301 CAA 303
          |||
Db   2956 CAA 2958


RESULT 7
AAC76445
ID   AAC76445 standard; cDNA; 2546 BP.
XX
AC   AAC76445;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human ORFX ORF2000 polynucleotide sequence SEQ ID NO:3999.
XX
KW   Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW   vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW   anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW   immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW   hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW   antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW   antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW   neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW   cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW   cholesterol ester storage; systemic lupus erythematosus; infection;
KW   severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW   allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW   bone damage; cartilage damage; antiinflammatory disease; coagulation;
```

```
KW   thrombosis; contraceptive; ss.
XX
OS   Homo sapiens.
XX
PN   WO200058473-A2.
XX
PD   05-OCT-2000.
XX
PF   31-MAR-2000; 2000WO-US08621.
XX
PR   31-MAR-1999;   99US-0127607.
PR   02-APR-1999;   99US-0127636.
PR   05-APR-1999;   99US-0127728.
PR   30-MAR-2000; 2000US-0540763.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2000-602362/57.
DR   P-PSDB; AAB42236.
XX
PT   Novel nucleic acids and peptides derived from open reading frame X,
PT   useful for treating e.g. cancers, proliferative disorders,
PT   neurodegenerative disorders and cardiovascular disease -
XX
PS   Claim 5; Page 3179-3180; 5507pp; English.
XX
CC   AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC   which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC   sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC   antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC   osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC   immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC   antidiabetic; hypotensive; dermatological; immunosuppressive;
CC   antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC   antithyroid; and antianaemic. The sequences can be used for determining
CC   the presence of or predisposition to, or preventing or treating
CC   pathological conditions associated with an ORFX-associated disorder. The
CC   nucleic acids can be used to express ORFX proteins in gene therapy
CC   vectors. The proteins and nucleic acids may be used to treat cancers,
CC   proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC   graft vs host disease, cardiovascular disease, diabetes mellitus,
CC   hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC   erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC   bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC   allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC   nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC   coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ   Sequence 2546 BP; 652 A; 643 C; 686 G; 564 T; 1 other;

  Query Match             59.8%;  Score 299.8;  DB 21;  Length 2546;
  Best Local Similarity   99.3%;  Pred. No. 8.1e-90;
  Matches  301;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2205 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTT 2264

Qy      61 AGGACTGGTGCCTGGCACAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2265 AGGACTGGTGCCTGGCACAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 2324

Qy     121 ACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCTTGGAAACAGACGGCGGCACCTTTC 180
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    2325 ACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCATGGAAACAGACGGCGGCACCTTTC 2384

Qy     181 CTCTAATCCAGCAAAGTGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTG 240
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    2385 CTCTAATCCAGCAAAATGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTG 2444

Qy     241 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2445 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 2504

Qy     301 CAA 303
          |||
Db    2505 CAA 2507


RESULT 8
AAS99132
ID   AAS99132 standard; cDNA; 2908 BP.
XX
AC   AAS99132;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Chimpanzee ELAC2 cDNA.
XX
KW   Human; mouse; HPC2; prostate cancer; neoplastic growth; cytostatic; ss;
KW   gene therapy; prostate cancer predisposing gene; chimpanzee; gorilla;
KW   sequencing primer; PCR primer.
XX
OS   Pan troglodytes.
XX
PN   WO200185911-A2.
XX
PD   15-NOV-2001.
XX
PF   07-MAY-2001; 2001WO-US14602.
XX
PR   05-MAY-2000; 2000US-0564805.
XX
PA   (MYRI-) MYRIAD GENETICS INC.
PA   (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI   Tavtigian SV,  Teng DHF,  Simard J,  Rommens JM;
XX
DR   WPI; 2002-066599/09.
DR   P-PSDB; AAU73592.
XX
PT   Novel nucleic acid sequence encoding HPC2 polypeptide, which is marker
PT   for prostate cancer, is useful in gene therapy techniques to restore
PT   HPC2 normal levels by which neoplastic growth is suppressed in
PT   recipient cell -
XX
PS   Claim 87; Page 198-201; 239pp; English.
XX
CC   The invention relates to a human prostate cancer predisposing gene coding
CC   for an HPC2 polypeptide. The DNA and protein sequences are useful as
CC   diagnostic reagents for identifying a mutant HPC2 nucleotide sequence in
CC   a suspected mutant HPC2 allele by comparing the sequence of the suspected
CC   mutant HPC2 allele with a wild-type HPC2 sequence. The sequences are also
CC   useful for detecting an alteration in HPC2, where the alteration is
CC   associated with cancer in a human. The method involves analysing an HPC2
CC   gene or an HPC2 gene expression product from a tissue of the human. The
CC   HPC2 gene is useful as a marker for prostate cancer and can be used in
CC   gene therapy techniques to suppress neoplastic growth of recipient cells
CC   which carry the mutant HPC2 allele. The sequences represent primers used
CC   in the methods of the invention, cDNA encoding human and mouse HPC2 and
CC   cDNA encoding HPC2 paralogues and orthologues.
XX
SQ   Sequence 2908 BP; 712 A; 788 C; 819 G; 589 T; 0 other;

  Query Match             59.2%;  Score 296.6;  DB 24;  Length 2908;
  Best Local Similarity   98.7%;  Pred. No. 1e-88;
  Matches  299;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTT 60
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    2606 GGTATGGAGCTGTGCCAAGGCTTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTT 2665

Qy      61 AGGACTGGTGCCTGGCACAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
```

```
Db    2666 AGGACTGGTGCCTGGCACAGCCGCGGGACAGGAGGCTGCCACACGGAAGCAAGCAGATGA 2725

Qy     121 ACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCTTGGAAACAGACGGCGGCACCTTTC 180
           |||||||||||||||||||||||||||||||| ||||||||||||||| |||||||||
Db    2726 ACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGGCTTGGAAACAGACGGCAGCACCTTTC 2785

Qy     181 CTCTAATCCAGCAAAGTGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2786 CTCTAATCCAGCAAAGTGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTG 2845

Qy     241 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2846 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 2905

Qy     301 CAA 303
           |||
Db    2906 CAA 2908
```

RESULT 9
ABN59829
ID ABN59829 standard; cDNA; 2992 BP.
XX
AC ABN59829;
XX
DT 28-JUN-2002 (first entry)
XX
DE Novel human coding sequence SEQ ID NO: 240.
XX
KW Human; antianaemic; vulnerary; antiinflammatory; immunomodulator;
KW antiinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
KW neuroprotective; antiparkinsonian; protein therapy; EST;
KW expressed sequence tag; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200222660-A2.
XX
PD 21-MAR-2002.
XX
PF 10-SEP-2001; 2001WO-US26015.
XX
PR 11-SEP-2000; 2000US-0659671.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
DR WPI; 2002-292408/33.
DR P-PSDB; ABB97416.
XX
PT An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis -
XX
PS Claim 1; SEQ ID NO 240; 509pp; English.
XX
CC The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a coding sequence of the
CC invention.
XX
SQ Sequence 2992 BP; 725 A; 807 C; 859 G; 601 T; 0 other;

```
  Query Match            59.2%;  Score 296.6;  DB 24;  Length 2992;
  Best Local Similarity  98.7%;  Pred. No. 1.1e-88;
  Matches  299;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTT 60
           |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    2680 GGTATGGAGCTGTGCCGAGGCTTGGGGTCCCACATAAGCACTAGTCTATAGATGCCTCTT 2739

Qy      61 AGGACTGGTGCCTGGCACAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
           |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    2740 AGGACTGGTGCCTGGCACAGCTGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 2799

Qy     121 ACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCTTGGAAACAGACGGCGGCACCTTTC 180
           |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    2800 ACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCATGGAAACAGACGGCGGCACCTTTC 2859

Qy     181 CTCTAATCCAGCAAAGTGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTG 240
           ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    2860 CTCTAATCCAGCAAAATGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTG 2919

Qy     241 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2920 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 2979

Qy     301 CAA 303
           |||
Db    2980 CAA 2982
```

RESULT 10
AAH14250
ID AAH14250 standard; cDNA; 2976 BP.
XX
AC AAH14250;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:11557.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 11557; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

```
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC   AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 2976 BP; 712 A; 807 C; 856 G; 601 T; 0 other;

  Query Match           58.8%;  Score 294.6;  DB 22;  Length 2976;
  Best Local Similarity  98.7%;  Pred. No. 5e-88;
  Matches  297;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTT 60
          ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db   2676 GGTATGGAGCTGTGCCGAGGCTTGGGGTCCCACATAAGCACTAGTCTATAGATGCCTCTT 2735

Qy     61 AGGACTGGTGCCTGGCACAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   2736 AGGACTGGTGCCTGGCACAGCTGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 2795

Qy    121 ACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCTTGGAAACAGACGGCGGCACCTTTC 180
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   2796 ACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCATGGAAACAGACGGCGGCACCTTTC 2855

Qy    181 CTCTAATCCAGCAAAGTGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTG 240
          ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   2856 CTCTAATCCAGCAAAATGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTG 2915

Qy    241 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2916 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 2975

Qy    301 C 301
          |
Db   2976 C 2976


RESULT 11
AAH10926/c
ID   AAH10926 standard; cDNA; 481 BP.
XX
AC   AAH10926;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA clone (3'-primer) SEQ ID NO:7761.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;   99JP-0248036.
PR   27-AUG-1999;   99JP-0300253.
PR   11-JAN-2000; 2000JP-0118776.
PR   02-MAY-2000; 2000JP-0183767.
PR   09-JUN-2000; 2000JP-0241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602
PT   full-length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full-length cDNAs -
XX
PS   Claim 3; SEQ ID 7761; 2537pp + CD ROM; English.
XX
CC   The present invention describes primer sets for synthesising 5602
CC   full-length cDNAs defined in the specification. Where a primer set
CC   comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC   to the complementary strand of a polynucleotide which comprises one of
CC   the 5602 nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC   AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 481 BP; 99 A; 130 C; 114 G; 131 T; 7 other;

  Query Match           58.6%;  Score 293.6;  DB 22;  Length 481;
  Best Local Similarity  98.3%;  Pred. No. 4e-88;
  Matches  296;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTT 60
          |||||||||||||||||||||||| |||||||||||||||||||||||||||||| ||||
Db    301 GGTATGGAGCTGTGCCGAGGCTTGGGGTCCCACATAAGCACTAGTCTATAGATGCNTCTT 242

Qy     61 AGGACTGGTGCCTGGCACAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    241 AGGACTGGTGCCTGGCACAGCTGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 182

Qy    121 ACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCTTGGAAACAGACGGCGGCACCTTTC 180
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    181 ACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCATGGAAACAGACGGCGGCACCTTTC 122

Qy    181 CTCTAATCCAGCAAAGTGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTG 240
          ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db    121 CTCTAATCCAGCAAAATGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTG 62

Qy    241 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 2

Qy    301 C 301
          |
Db      1 C 1


RESULT 12
AAS99133
ID   AAS99133 standard; cDNA; 2892 BP.
XX
```

```
AC   AAS99133;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Gorilla ELAC2 cDNA.
XX
KW   Human; mouse; HPC2; prostate cancer; neoplastic growth; cytostatic; ss;
KW   gene therapy; prostate cancer predisposing gene; chimpanzee; gorilla;
KW   sequencing primer; PCR primer.
XX
OS   Gorilla gorilla.
XX
PN   WO200185911-A2.
XX
PD   15-NOV-2001.
XX
PF   07-MAY-2001; 2001WO-US14602.
XX
PR   05-MAY-2000; 2000US-0564805.
XX
PA   (MYRI-) MYRIAD GENETICS INC.
PA   (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI   Tavtigian SV,  Teng DHF,  Simard J,  Rommens JM;
XX
DR   WPI; 2002-066599/09.
DR   P-PSDB; AAU73593.
XX
PT   Novel nucleic acid sequence encoding HPC2 polypeptide, which is marker
PT   for prostate cancer, is useful in gene therapy techniques to restore
PT   HPC2 normal levels by which neoplastic growth is suppressed in
PT   recipient cell  -
XX
PS   Claim 92; Page 204-207; 239pp; English.
XX
CC   The invention relates to a human prostate cancer predisposing gene coding
CC   for an HPC2 polypeptide. The DNA and protein sequences are useful as
CC   diagnostic reagents for identifying a mutant HPC2 nucleotide sequence in
CC   a suspected mutant HPC2 allele by comparing the sequence of the suspected
CC   mutant HPC2 allele with a wild-type HPC2 sequence. The sequences are also
CC   useful for detecting an alteration in HPC2, where the alteration is
CC   associated with cancer in a human. The method involves analysing an HPC2
CC   gene or an HPC2 gene expression product from a tissue of the human. The
CC   HPC2 gene is useful as a marker for prostate cancer and can be used in
CC   gene therapy techniques to suppress neoplastic growth of recipient cells
CC   which carry the mutant HPC2 allele. The sequences represent primers used
CC   in the methods of the invention, cDNA encoding human and mouse HPC2 and
CC   cDNA encoding HPC2 paralogues and orthologues.
XX
SQ   Sequence 2892 BP; 704 A; 787 C; 815 G; 586 T; 0 other;

  Query Match            50.8%;  Score 254.6;  DB 24;  Length 2892;
  Best Local Similarity  93.4%;  Pred. No. 1.4e-74;
  Matches  283;  Conservative    0;  Mismatches    4;  Indels   16;  Gaps    1;

Qy      1 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTT 60
          |||||||||||||||||||||| |||||||||||||||||||||||||||
Db   2606 GGTATGGAGCTGTGCCGAGGCTTAGGCTCCCACATAAGCACTAGTCTATA---------- 2655

Qy     61 AGGACTGGTGCCTGGCACAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   2656 ------GGTGCCTGGCACAGCCGCGGGACAGGAGGCTGCCACACGGAAGCAAGCAGATGA 2709

Qy    121 ACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCTTGGAAACAGACGGCGGCACCTTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   2710 ACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCTTGGAAACAGACGGCAGCACCTTTC 2769

Qy    181 CTCTAATCCAGCAAAGTGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTG 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db   2770 CTCTAATCCAGCAAAGTGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCACTG 2829

Qy    241 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   2830 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 2889

Qy    301 CAA 303
          |||
Db   2890 CAA 2892


RESULT 13
AAS72207
ID   AAS72207 standard; cDNA; 1402 BP.
XX
AC   AAS72207;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #8011.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG08020.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 8011; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1402 BP; 338 A; 371 C; 377 G; 316 T; 0 other;

  Query Match            40.7%;  Score 204;  DB 23;  Length 1402;
  Best Local Similarity  97.6%;  Pred. No. 9.6e-58;
  Matches  207;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;
```

```
Qy      91 GGAGGCTGCCACACGGAAGCAAGCAGATGAACTAATTTCATTTCAAGGCAGTTTTTAAAG 150
           | | | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1191 GAAAGATGCCACACGGAAGCAAGCAGATGAACTAATTTCATTTCAAGGCAGTTTTTAAAG 1250

Qy     151 AAGTCTTGGAAACAGACGGCGGCACCTTTCCTCTAATCCAGCAAAGTGATTCCCTGCACA 210
           ||||| ||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    1251 AAGTCATGGAAACAGACGGCGGCACCTTTCCTCTAATCCAGCAAAATGATTCCCTGCACA 1310

Qy     211 CCAGAGACAAGCAGAGTAACAGGATCAGTGGGTCTAAGTGTCCGAGACTTAACGAAAATA 270
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1311 CCAGAGACAAGCAGAGTAACAGGATCAGTGGGTCTAAGTGTCCGAGACTTAACGAAAATA 1370

Qy     271 GTATTTCAGCTGCAATAAAGATTGAGTTTGCA 302
           ||||||||||||||||||||||||||||||||
Db    1371 GTATTTCAGCTGCAATAAAGATTGAGTTTGCA 1402
```

RESULT 14
AAF64286
ID AAF64286 standard; cDNA; 386 BP.
XX
AC AAF64286;
XX
DT 09-APR-2001 (first entry)
XX
DE Novel human polynucleotide, SEQ ID NO: 42.
XX
KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW breast cancer; lung cancer; cancer detection; ss.
XX
OS Homo sapiens.
XX
PN WO200102568-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18374.
XX
PR 02-JUL-1999; 99US-0142310.
PR 02-JUL-1999; 99US-0142311.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
PI Crkenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitiz D;
PI Kita D, Garcia V, Jones LW, Strache-Crain B;
XX
DR WPI; 2001-091805/10.
XX
PT Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
XX
PS Claim 9; Page 550-551; 1046pp; English.
XX
CC The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.
XX

SQ Sequence 386 BP; 82 A; 90 C; 111 G; 103 T; 0 other;

Query Match 31.5%; Score 158; DB 22; Length 386;
Best Local Similarity 100.0%; Pred. No. 1.4e-42;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     344 AGGGTCTGCTGTGCACCACCTTGGAGAAGGCTCTCTGTGCTGTAGTGTGGCAGCTGCCTG 403
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7 AGGGTCTGCTGTGCACCACCTTGGAGAAGGCTCTCTGTGCTGTAGTGTGGCAGCTGCCTG 66

Qy     404 GTACCCGGGTGGCTTGGAAGAAGTCAGCTCCCGTCGTAGTGAGCACCTCTGGAACCTGTC 463
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67 GTACCCGGGTGGCTTGGAAGAAGTCAGCTCCCGTCGTAGTGAGCACCTCTGGAACCTGTC 126

Qy     464 CTCAGAGAGCCACCCTTATTCGCCAAGTCTTTTTGACA 501
           |||||||||||||||||||||||||||||||||||||||
Db     127 CTCAGAGAGCCACCCTTATTCGCCAAGTCTTTTTGACA 164
```

RESULT 15
AAS72208
ID AAS72208 standard; cDNA; 1450 BP.
XX
AC AAS72208;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #8012.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG08021.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 8012; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and

```
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1450 BP; 355 A; 382 C; 418 G; 294 T; 1 other;

  Query Match             30.9%;  Score 155;  DB 23;  Length 1450;
  Best Local Similarity   87.8%;  Pred. No. 2.9e-41;
  Matches  273;  Conservative    0;  Mismatches   25;  Indels   13;  Gaps    9;

Qy       6 GGAGCTGTGCCGAGGC--TTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTTAGG 63
           | ||||||||||||||  | |||  ||||||||||||||||||||||||||||||   |  |
Db    1026 GAAGCTGTGCCGAGGCCTTGGGGTCCCCACATAAGCACTAGTCTATAGATGCCCTCTTAG 1085

Qy      64 ---ACTGGTGCCTGGCACAGCCGCGGGCCAGGAGGCTG-CCACACGGAAGCAAG-CAGAT 118
               |   |||||||||| |||| |||||||||| |||||||||||||| ||||| 
Db    1086 GAACTGGTGCCCTGGCACAGCTGCGGCCCAGGAGGCTGCCCACACGGAAGCAAGCCAGAT 1145

Qy     119 GAACTAATTTCATTT-CAAGGCAGTTTTTAAAG-AAGTCTTGGAAACAGACGGCGGCACC 176
           ||||||||||||||| |||||||||||||||| ||||| ||||||||||||||||||||
Db    1146 GAACTAATTTCATTTCCAAGGCAGTTTTTAAAGAAAGTCATGGAAACAGACGGCGGCACC 1205

Qy     177 TTTCCTCTAATCCAGCAAAGTGATTCCCTGCACACCAGAGACAAGCAGA-GTAACAGGAT 235
           |||||||||||||||||| |||||||||||||||||||||||||||||| ||||||||||
Db    1206 TTTCCTCTAATCCAGCAAAATGATTCCCTGCACACCAGAGACAAGCAGAGGTAACAGGAT 1265

Qy     236 CA--GTGGGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAG-CTGCAATAAAGAT 292
           ||    ||    |||||||||||||||||||||||||||||||||| ||||||||||||
Db    1266 CAAGTGGGTCTAAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGTCTGCAATAAAGAT 1325

Qy     293 TGAGTTTGCAA 303
           ||||||||||| 
Db    1326 TGAGTTTGCAA 1336


Search completed: May 16, 2003, 03:40:57
Job time : 188.593 secs
```

GenCore version 5.1.5

OM nucleic - nucleic search, using sw model

Run on: May 16, 2003, 01:11:33 ; Search time 552.621 Seconds
(without alignments)
10110.387 Million cell updates/sec

Title: US-09-434-382-1
Perfect score: 2481
Sequence: 1 atgtgggcgctttgctcgct.....agaaggtcagagcccagtga 2481

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2481	100.0	2481	24	AAS98916	Human prostate can
	2	2481	100.0	2958	21	AAA58453	Human prostate can
	3	2481	100.0	2958	24	AAS98917	Human prostate can
	4	2481	100.0	2992	24	ABN59829	Novel human coding
	5	2471.6	99.6	2478	21	AAA52810	Human sulphatase G
	6	2455.4	99.0	2908	24	AAS99132	Chimpanzee ELAC2 c
	7	2448.2	98.7	2976	22	AAH14250	Human cDNA sequenc
	8	2442.6	98.5	2892	24	AAS99133	Gorilla ELAC2 cDNA
	9	1782	71.8	2546	21	AAC76445	Human ORFX ORF2000
	10	1645.6	66.3	2470	24	AAS99131	Mouse ELAC2 cDNA.
	11	1080.8	43.6	1402	23	AAS72207	DNA encoding novel
	12	518.2	20.9	1450	23	AAS72208	DNA encoding novel
	13	475.8	19.2	584	22	AAH05835	Human cDNA clone (
	14	247.4	10.0	350	24	AAS99124	Human prostate can
	15	247.4	10.0	26664	21	AAA60207	Human prostate can
	16	247.4	10.0	26664	24	AAS98942	Human prostate can
c	17	247.4	10.0	37959	22	AAK81763	Human immune/haema
	18	245	9.9	295	24	AAS98918	Human prostate can
	19	237	9.6	238	21	AAZ80231	Human colon cancer
	20	228	9.2	655	24	AAS98941	Human prostate can
	21	212.4	8.6	367	23	AAS72206	DNA encoding novel
	22	210.2	8.5	290	24	ABN18966	Human ORFX polynuc
c	23	165.8	6.7	622	24	ABQ15916	Oligonucleotide fo
	24	165.8	6.7	622	24	ABQ15917	Oligonucleotide fo
	25	165.4	6.7	622	24	ABQ15914	Oligonucleotide fo
c	26	165.4	6.7	622	24	ABQ15915	Oligonucleotide fo
	27	145	5.8	145	24	AAS98940	Human prostate can
	28	139	5.6	139	24	AAS98930	Human prostate can
	29	139	5.6	139	24	AAS98934	Human prostate can
	30	133.4	5.4	2232	23	ABL02765	Drosophila melanog
	31	133.4	5.4	4290	23	ABL02764	Drosophila melanog
	32	121	4.9	121	24	AAS98938	Human prostate can
	33	120	4.8	120	24	AAS98924	Human prostate can
	34	119	4.8	119	24	AAS98932	Human prostate can
	35	113.6	4.6	326	21	AAA60390	Murine prostate ca
	36	113.6	4.6	326	24	AAS99125	Mouse prostate can
	37	113	4.6	113	24	AAS98928	Human prostate can
	38	110	4.4	110	24	AAS98936	Human prostate can
	39	100	4.0	100	24	AAS98937	Human prostate can
	40	99.2	4.0	122	22	AAL07785	Human breast cance
	41	99	4.0	107	22	AAL16688	Human breast cance
c	42	98	4.0	274	22	AAL16514	Human breast cance
	43	97	3.9	97	24	AAS98933	Human prostate can
	44	96	3.9	96	24	AAS98929	Human prostate can
c	45	96	3.9	260	22	AAL25357	Human breast cance

ALIGNMENTS

RESULT 1
AAS98916
ID AAS98916 standard; cDNA; 2481 BP.
XX
AC AAS98916;
XX
DT 12-MAR-2002 (first entry)
XX
DE Human prostate cancer predisposing gene (HPC2) cDNA coding sequence.
XX
KW Human; mouse; HPC2; prostate cancer; neoplastic growth; cytostatic; ss;
KW gene therapy; prostate cancer predisposing gene; chimpanzee; gorilla;
KW sequencing primer; PCR primer.
XX
OS Homo sapiens.
XX
PN WO200185911-A2.
XX
PD 15-NOV-2001.
XX
PF 07-MAY-2001; 2001WO-US14602.
XX
PR 05-MAY-2000; 2000US-0564805.
XX
PA (MYRI-) MYRIAD GENETICS INC.
PA (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI Tavtigian SV, Teng DHF, Simard J, Rommens JM;
XX
DR WPI; 2002-066599/09.
DR P-PSDB; AAU73586.
XX

PT Novel nucleic acid sequence encoding HPC2 polypeptide, which is marker
PT for prostate cancer, is useful in gene therapy techniques to restore
PT HPC2 normal levels by which neoplastic growth is suppressed in
PT recipient cell -
XX
PS Claim 2; Page 128-131; 239pp; English.
XX
CC The invention relates to a human prostate cancer predisposing gene coding
CC for an HPC2 polypeptide. The DNA and protein sequences are useful as
CC diagnostic reagents for identifying a mutant HPC2 nucleotide sequence in
CC a suspected mutant HPC2 allele by comparing the sequence of the suspected
CC mutant HPC2 allele with a wild-type HPC2 sequence. The sequences are also
CC useful for detecting an alteration in HPC2, where the alteration is
CC associated with cancer in a human. The method involves analysing an HPC2
CC gene or an HPC2 gene expression product from a tissue of the human. The
CC HPC2 gene is useful as a marker for prostate cancer and can be used in
CC gene therapy techniques to suppress neoplastic growth of recipient cells
CC which carry the mutant HPC2 allele. The sequences represent primers used
CC in the methods of the invention, cDNA encoding human and mouse HPC2 and
CC cDNA encoding HPC2 paralogues and orthologues.
XX
SQ Sequence 2481 BP; 588 A; 686 C; 711 G; 496 T; 0 other;

Query Match 100.0%; Score 2481; DB 24; Length 2481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 ATGTGGGCGCTTTGCTCGCTGCTGCGGTCCGCGGCCGGACGCACCATGTCGCAGGGACGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGTGGGCGCTTTGCTCGCTGCTGCGGTCCGCGGCCGGACGCACCATGTCGCAGGGACGC 60

Qy     61 ACCATATCGCAGGCACCCGCCCGCCGCGAGCGGCCGCGCAAGGACCCGCTGCGGCACCTG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ACCATATCGCAGGCACCCGCCCGCCGCGAGCGGCCGCGCAAGGACCCGCTGCGGCACCTG 120

Qy    121 CGCACGCGAGAGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAAACACCGTGTACCTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CGCACGCGAGAGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAAACACCGTGTACCTG 180

Qy    181 CAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTC 240

Qy    241 AACCGGTATCTCTTCAACTGTGGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 AACCGGTATCTCTTCAACTGTGGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTA 300

Qy    301 AAGGTTGCTCGCCTGGACAACATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AAGGTTGCTCGCCTGGACAACATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGC 360

Qy    361 TTAAGTGGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTAAGTGGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 420

Qy    421 CCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGA 480

Qy    481 ATAGAACTGGCTGTGCGGCCCCACTCTGCCCCAGAATACGAGGATGAAACCATGACAGTT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ATAGAACTGGCTGTGCGGCCCCACTCTGCCCCAGAATACGAGGATGAAACCATGACAGTT 540

Qy    541 TACCAGATCCCCATACACAGTGAACAGAGGAGGGGAAAGCACCAACCATGGCAGAGTCCA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TACCAGATCCCCATACACAGTGAACAGAGGAGGGGAAAGCACCAACCATGGCAGAGTCCA 600

Qy    601 GAAAGGCCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCGAATGAAAAT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GAAAGGCCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCGAATGAAAAT 660

Qy    661 GAGCCACACCTTCCACATGGTGTTAGCCAGAGAAGAGGGGTCAGGGACTCTTCCCTGGTC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GAGCCACACCTTCCACATGGTGTTAGCCAGAGAAGAGGGGTCAGGGACTCTTCCCTGGTC 720

Qy    721 GTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGGAAACTTCTTGGTGCTCAAAGCAAAG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGGAAACTTCTTGGTGCTCAAAGCAAAG 780

Qy    781 GAGATGGGCCTCCCAGTTGGGACAGCTGCCATCGCTCCCATCATTGCTGCTGTCAAGGAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GAGATGGGCCTCCCAGTTGGGACAGCTGCCATCGCTCCCATCATTGCTGCTGTCAAGGAC 840

Qy    841 GGGAAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCCTCCA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 GGGAAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCCTCCA 900

Qy    901 GATCCTGGTGCTGCTTTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCATTCAACCCATC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GATCCTGGTGCTGCTTTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCATTCAACCCATC 960

Qy    961 TGTGAGAATGCCACCTTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCGTGGCCTTGGTG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TGTGAGAATGCCACCTTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCGTGGCCTTGGTG 1020

Qy   1021 GTTCACATGGCCCCAGCATCTGTGCTTGTGGACAGCAGGTACCAGCAGTGGATGGAGAGG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GTTCACATGGCCCCAGCATCTGTGCTTGTGGACAGCAGGTACCAGCAGTGGATGGAGAGG 1080

Qy   1081 TTTGGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TTTGGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTT 1140

Qy   1141 CGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCTC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCTC 1200

Qy   1201 ACCAGTTTCCGCTGTAAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGGTTCAGGGTGAA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ACCAGTTTCCGCTGTAAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGGTTCAGGGTGAA 1260

Qy   1261 TGCCTCCTCAAGTACCAGCTCCGTCCCAGGAGGGAGTGGCAGAGGGATGCCATTATTACT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 TGCCTCCTCAAGTACCAGCTCCGTCCCAGGAGGGAGTGGCAGAGGGATGCCATTATTACT 1320

Qy   1321 TGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGCGTG 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 TGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGCGTG 1380

Qy   1381 CAGGAGTACAGGAGGAGTGCGCAGGACGGCCCAGCCCCAGCAGAGAAAAGAAGTCAGTAC 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CAGGAGTACAGGAGGAGTGCGCAGGACGGCCCAGCCCCAGCAGAGAAAAGAAGTCAGTAC 1440

Qy   1441 CCAGAAATCATCTTCCTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGT 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 CCAGAAATCATCTTCCTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGT 1500

Qy   1501 GCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACA 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACA 1560

Qy   1561 TTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 TTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCT 1620

Qy   1621 GCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGGGCTTGCCAAGTATCTTGCTG 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGGGCTTGCCAAGTATCTTGCTG 1680

Qy   1681 CAGAGAGAACGCGCCTTGGCATCTTTGGGAAAGCCGCTTCACCCTTTGCTGGTGGTTGCC 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 CAGAGAGAACGCGCCTTGGCATCTTTGGGAAAGCCGCTTCACCCTTTGCTGGTGGTTGCC 1740

Qy   1741 CCCAACCAGCTCAAAGCCTGGCTCCAGCAGTACCACAACCAGTGCCAGGAGGTCCTGCAC 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 CCCAACCAGCTCAAAGCCTGGCTCCAGCAGTACCACAACCAGTGCCAGGAGGTCCTGCAC 1800
```

```
Qy   1801 CACATCAGTATGATTCCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCTCCAGTCCTGCA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 CACATCAGTATGATTCCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCTCCAGTCCTGCA 1860

Qy   1861 GTGGAAAGATTGATCAGTTCGCTGTTGCGAACATGTGATTTGGAAGAGTTTCAGACCTGT 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 GTGGAAAGATTGATCAGTTCGCTGTTGCGAACATGTGATTTGGAAGAGTTTCAGACCTGT 1920

Qy   1921 CTGGTGCGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAA 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CTGGTGCGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAA 1980

Qy   1981 GTGGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCTGGTCCGGATGGGGAAAGATGCC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 GTGGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCTGGTCCGGATGGGGAAAGATGCC 2040

Qy   2041 ACCCTCCTGATACATGAAGCCACCCTGGAAGATGGTTTGGAAGAGGAAGCAGTGGAAAAG 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 ACCCTCCTGATACATGAAGCCACCCTGGAAGATGGTTTGGAAGAGGAAGCAGTGGAAAAG 2100

Qy   2101 ACACACAGCACAACGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTCATT 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 ACACACAGCACAACGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTCATT 2160

Qy   2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCCCTCTTCAGCCCCAACTTCAGC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCCCTCTTCAGCCCCAACTTCAGC 2220

Qy   2221 GAGAAAGTGGGAGTTGCCTTTGACCACATGAAGGTCTGCTTTGGAGACTTTCCAACAATG 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 GAGAAAGTGGGAGTTGCCTTTGACCACATGAAGGTCTGCTTTGGAGACTTTCCAACAATG 2280

Qy   2281 CCCAAGCTGATTCCCCCACTGAAAGCCCTGTTTGCTGGCGACATCGAGGAGATGGAGGAG 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 CCCAAGCTGATTCCCCCACTGAAAGCCCTGTTTGCTGGCGACATCGAGGAGATGGAGGAG 2340

Qy   2341 CGCAGGGAGAAGCGGGAGCTGCGGCAGGTGCGGGCGGCCCTCCTGTCCAGGGAGCTGGCA 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 CGCAGGGAGAAGCGGGAGCTGCGGCAGGTGCGGGCGGCCCTCCTGTCCAGGGAGCTGGCA 2400

Qy   2401 GGCGGCCTGGAGGATGGGGAGCCTCAGCAGAAGCGGGCCCACACAGAGGAGCCACAGGCC 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 GGCGGCCTGGAGGATGGGGAGCCTCAGCAGAAGCGGGCCCACACAGAGGAGCCACAGGCC 2460

Qy   2461 AAGAAGGTCAGAGCCCAGTGA 2481
          |||||||||||||||||||||
Db   2461 AAGAAGGTCAGAGCCCAGTGA 2481
```

```
RESULT 2
AAA58453
ID   AAA58453 standard; cDNA; 2958 BP.
XX
AC   AAA58453;
XX
DT   07-DEC-2000  (first entry)
XX
DE   Human prostate cancer predisposing gene HPC2 coding sequence.
XX
KW   Human; prostate cancer predisposing gene; HPC2; chromosome 17p;
KW   gene therapy; peptide therapy; drug design; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             51..2531
FT                   /*tag= a
FT                   /product= "HPC2"
XX
PN   WO200027864-A1.
XX
PD   18-MAY-2000.
XX
PF   05-NOV-1999;   99WO-US26055.
XX
PR   06-NOV-1998;   98US-0107468.
XX
PA   (MYRI-) MYRIAD GENETICS INC.
XX
PI   Tavtigian SV,  Teng DHF,  Simard J,  Rommens JM;
XX
DR   WPI; 2000-376481/32.
DR   P-PSDB; AAB07228.
XX
PT   Human prostate cancer (HPC)2 nucleic acids, polypeptides, and
PT   antibodies, useful for treatment and diagnosis of prostate cancer -
XX
PS   Claim 3; Page 98-100; 157pp; English.
XX
CC   The present sequence is the coding sequence of the human prostate
CC   cancer predisposing gene HPC2, which is found on chromosome 17p. Some
CC   alleles of this gene cause a predisposition to cancer, particularly
CC   prostate cancer. This gene and its protein can be used in peptide and
CC   gene therapy for cancer patients, as well as being useful as diagnostic
CC   tools (both for cancer sufferers and those with a predisposition to the
CC   disease) and in the production of cancer drugs. This sequence was
CC   isolated by cloning and sequencing the region of the genome which
CC   appeared to cause a predisposition to prostate cancer.
XX
SQ   Sequence 2958 BP; 707 A; 805 C; 848 G; 598 T; 0 other;
```

```
  Query Match             100.0%;  Score 2481;  DB 21;  Length 2958;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2481;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ATGTGGGCGCTTTGCTCGCTGCTGCGGTCCGCGGCCGGACGCACCATGTCGCAGGGACGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     51 ATGTGGGCGCTTTGCTCGCTGCTGCGGTCCGCGGCCGGACGCACCATGTCGCAGGGACGC 110

Qy     61 ACCATATCGCAGGCACCCGCCCGCCGCGAGCGGCCGCGCAAGGACCCGCTGCGGCACCTG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    111 ACCATATCGCAGGCACCCGCCCGCCGCGAGCGGCCGCGCAAGGACCCGCTGCGGCACCTG 170

Qy    121 CGCACGCGAGAGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAAACACCGTGTACCTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    171 CGCACGCGAGAGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAAACACCGTGTACCTG 230

Qy    181 CAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    231 CAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTC 290

Qy    241 AACCGGTATCTCTTCAACTGTGGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    291 AACCGGTATCTCTTCAACTGTGGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTA 350

Qy    301 AAGGTTGCTCGCCTGGACAACATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    351 AAGGTTGCTCGCCTGGACAACATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGC 410

Qy    361 TTAAGTGGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    411 TTAAGTGGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 470

Qy    421 CCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    471 CCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGA 530

Qy    481 ATAGAACTGGCTGTGCGGCCCCACTCTGCCCCAGAATACGAGGATGAAACCATGACAGTT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    531 ATAGAACTGGCTGTGCGGCCCCACTCTGCCCCAGAATACGAGGATGAAACCATGACAGTT 590

Qy    541 TACCAGATCCCCATACACAGTGAACAGAGGAGGGGAAAGCACCAACCATGGCAGAGTCCA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    591 TACCAGATCCCCATACACAGTGAACAGAGGAGGGGAAAGCACCAACCATGGCAGAGTCCA 650
```

```
Qy    601 GAAAGGCCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCGAATGAAAAT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    651 GAAAGGCCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCGAATGAAAAT 710

Qy    661 GAGCCACACCTTCCACATGGTGTTAGCCAGAGAAGAGGGGTCAGGGACTCTTCCCTGGTC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    711 GAGCCACACCTTCCACATGGTGTTAGCCAGAGAAGAGGGGTCAGGGACTCTTCCCTGGTC 770

Qy    721 GTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGGAAACTTCTTGGTGCTCAAAGCAAAG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    771 GTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGGAAACTTCTTGGTGCTCAAAGCAAAG 830

Qy    781 GAGATGGGCCTCCCAGTTGGGACAGCTGCCATCGCTCCCATCATTGCTGCTGTCAAGGAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    831 GAGATGGGCCTCCCAGTTGGGACAGCTGCCATCGCTCCCATCATTGCTGCTGTCAAGGAC 890

Qy    841 GGGAAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCCTCCA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    891 GGGAAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCCTCCA 950

Qy    901 GATCCTGGTGCTGCTTTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCATTCAACCCATC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    951 GATCCTGGTGCTGCTTTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCATTCAACCCATC 1010

Qy    961 TGTGAGAATGCCACCTTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCGTGGCCTTGGTG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1011 TGTGAGAATGCCACCTTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCGTGGCCTTGGTG 1070

Qy   1021 GTTCACATGGCCCCAGCATCTGTGCTTGTGGACAGCAGGTACCAGCAGTGGATGGAGAGG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1071 GTTCACATGGCCCCAGCATCTGTGCTTGTGGACAGCAGGTACCAGCAGTGGATGGAGAGG 1130

Qy   1081 TTTGGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1131 TTTGGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTT 1190

Qy   1141 CGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCTC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1191 CGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCTC 1250

Qy   1201 ACCAGTTTCCGCTGTAAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGGTTCAGGGTGAA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1251 ACCAGTTTCCGCTGTAAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGGTTCAGGGTGAA 1310

Qy   1261 TGCCTCCTCAAGTACCAGCTCCGTCCCAGGAGGGAGTGGCAGAGGGATGCCATTATTACT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1311 TGCCTCCTCAAGTACCAGCTCCGTCCCAGGAGGGAGTGGCAGAGGGATGCCATTATTACT 1370

Qy   1321 TGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGCGTG 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1371 TGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGCGTG 1430

Qy   1381 CAGGAGTACAGGAGGAGTGCGCAGGACGGCCCAGCCCCAGCAGAGAAAAGAAGTCAGTAC 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1431 CAGGAGTACAGGAGGAGTGCGCAGGACGGCCCAGCCCCAGCAGAGAAAAGAAGTCAGTAC 1490

Qy   1441 CCAGAAATCATCTTCCTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGT 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1491 CCAGAAATCATCTTCCTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGT 1550

Qy   1501 GCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACA 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1551 GCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACA 1610

Qy   1561 TTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1611 TTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCT 1670

Qy   1621 GCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGGGCTTGCCAAGTATCTTGCTG 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1671 GCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGGGCTTGCCAAGTATCTTGCTG 1730

Qy   1681 CAGAGAGAACGCGCCTTGGCATCTTTGGGAAAGCCGCTTCACCCTTTGCTGGTGGTTGCC 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1731 CAGAGAGAACGCGCCTTGGCATCTTTGGGAAAGCCGCTTCACCCTTTGCTGGTGGTTGCC 1790

Qy   1741 CCCAACCAGCTCAAAGCCTGGCTCCAGCAGTACCACAACCAGTGCCAGGAGGTCCTGCAC 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1791 CCCAACCAGCTCAAAGCCTGGCTCCAGCAGTACCACAACCAGTGCCAGGAGGTCCTGCAC 1850

Qy   1801 CACATCAGTATGATTCCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCTCCAGTCCTGCA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1851 CACATCAGTATGATTCCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCTCCAGTCCTGCA 1910

Qy   1861 GTGGAAAGATTGATCAGTTCGCTGTTGCGAACATGTGATTTGGAAGAGTTTCAGACCTGT 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1911 GTGGAAAGATTGATCAGTTCGCTGTTGCGAACATGTGATTTGGAAGAGTTTCAGACCTGT 1970

Qy   1921 CTGGTGCGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAA 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1971 CTGGTGCGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAA 2030

Qy   1981 GTGGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCTGGTCCGGATGGGGAAAGATGCC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2031 GTGGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCTGGTCCGGATGGGGAAAGATGCC 2090

Qy   2041 ACCCTCCTGATACATGAAGCCACCCTGGAAGATGGTTTGGAAGAGGAAGCAGTGGAAAAG 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2091 ACCCTCCTGATACATGAAGCCACCCTGGAAGATGGTTTGGAAGAGGAAGCAGTGGAAAAG 2150

Qy   2101 ACACACAGCACAACGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTCATT 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2151 ACACACAGCACAACGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTCATT 2210

Qy   2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCCCTCTTCAGCCCCAACTTCAGC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2211 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCCCTCTTCAGCCCCAACTTCAGC 2270

Qy   2221 GAGAAAGTGGGAGTTGCCTTTGACCACATGAAGGTCTGCTTTGGAGACTTTCCAACAATG 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2271 GAGAAAGTGGGAGTTGCCTTTGACCACATGAAGGTCTGCTTTGGAGACTTTCCAACAATG 2330

Qy   2281 CCCAAGCTGATTCCCCCACTGAAAGCCCTGTTTGCTGGCGACATCGAGGAGATGGAGGAG 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2331 CCCAAGCTGATTCCCCCACTGAAAGCCCTGTTTGCTGGCGACATCGAGGAGATGGAGGAG 2390

Qy   2341 CGCAGGGAGAAGCGGGAGCTGCGGCAGGTGCGGGCGGCCCTCCTGTCCAGGGAGCTGGCA 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2391 CGCAGGGAGAAGCGGGAGCTGCGGCAGGTGCGGGCGGCCCTCCTGTCCAGGGAGCTGGCA 2450

Qy   2401 GGCGGCCTGGAGGATGGGGAGCCTCAGCAGAAGCGGGCCCACACAGAGGAGCCACAGGCC 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2451 GGCGGCCTGGAGGATGGGGAGCCTCAGCAGAAGCGGGCCCACACAGAGGAGCCACAGGCC 2510

Qy   2461 AAGAAGGTCAGAGCCCAGTGA 2481
          ||||||||||||||||||||
Db   2511 AAGAAGGTCAGAGCCCAGTGA 2531
```

RESULT 3
AAS98917
ID AAS98917 standard; cDNA; 2958 BP.
XX
AC AAS98917;
XX
DT 12-MAR-2002 (first entry)
XX
DE Human prostate cancer predisposing gene (HPC2) extended cDNA.
XX
KW Human; mouse; HPC2; prostate cancer; neoplastic growth; cytostatic; ss;
KW gene therapy; prostate cancer predisposing gene; chimpanzee; gorilla;
KW sequencing primer; PCR primer.
XX
OS Homo sapiens.
XX
PN WO200185911-A2.

XX
PD 15-NOV-2001.
XX
PF 07-MAY-2001; 2001WO-US14602.
XX
PR 05-MAY-2000; 2000US-0564805.
XX
PA (MYRI-) MYRIAD GENETICS INC.
PA (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI Tavtigian SV, Teng DHF, Simard J, Rommens JM;
XX
DR WPI; 2002-066599/09.
DR P-PSDB; AAU73586.
XX
PT Novel nucleic acid sequence encoding HPC2 polypeptide, which is marker
PT for prostate cancer, is useful in gene therapy techniques to restore
PT HPC2 normal levels by which neoplastic growth is suppressed in
PT recipient cell -
XX
PS Claim 3; Page 134-136; 239pp; English.
XX
CC The invention relates to a human prostate cancer predisposing gene coding
CC for an HPC2 polypeptide. The DNA and protein sequences are useful as
CC diagnostic reagents for identifying a mutant HPC2 nucleotide sequence in
CC a suspected mutant HPC2 allele by comparing the sequence of the suspected
CC mutant HPC2 allele with a wild-type HPC2 sequence. The sequences are also
CC useful for detecting an alteration in HPC2, where the alteration is
CC associated with cancer in a human. The method involves analysing an HPC2
CC gene or an HPC2 gene expression product from a tissue of the human. The
CC HPC2 gene is useful as a marker for prostate cancer and can be used in
CC gene therapy techniques to suppress neoplastic growth of recipient cells
CC which carry the mutant HPC2 allele. The sequences represent primers used
CC in the methods of the invention, cDNA encoding human and mouse HPC2 and
CC cDNA encoding HPC2 paralogues and orthologues.
XX
SQ Sequence 2958 BP; 707 A; 805 C; 848 G; 598 T; 0 other;

Query Match 100.0%; Score 2481; DB 24; Length 2958;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 ATGTGGGCGCTTTGCTCGCTGCTGCGGTCCGCGGCCGGACGCACCATGTCGCAGGGACGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51 ATGTGGGCGCTTTGCTCGCTGCTGCGGTCCGCGGCCGGACGCACCATGTCGCAGGGACGC 110

Qy      61 ACCATATCGCAGGCACCCGCCCGCCGCGAGCGGCCGCGCAAGGACCCGCTGCGGCACCTG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     111 ACCATATCGCAGGCACCCGCCCGCCGCGAGCGGCCGCGCAAGGACCCGCTGCGGCACCTG 170

Qy     121 CGCACGCGAGAGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAAACACCGTGTACCTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     171 CGCACGCGAGAGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAAACACCGTGTACCTG 230

Qy     181 CAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     231 CAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTC 290

Qy     241 AACCGGTATCTCTTCAACTGTGGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     291 AACCGGTATCTCTTCAACTGTGGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTA 350

Qy     301 AAGGTTGCTCGCCTGGACAACATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     351 AAGGTTGCTCGCCTGGACAACATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGC 410

Qy     361 TTAAGTGGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     411 TTAAGTGGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 470

Qy     421 CCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     471 CCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGA 530

Qy     481 ATAGAACTGGCTGTGCGGCCCCACTCTGCCCCAGAATACGAGGATGAAACCATGACAGTT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     531 ATAGAACTGGCTGTGCGGCCCCACTCTGCCCCAGAATACGAGGATGAAACCATGACAGTT 590

Qy     541 TACCAGATCCCCATACACAGTGAACAGAGGAGGGGAAAGCACCAACCATGGCAGAGTCCA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     591 TACCAGATCCCCATACACAGTGAACAGAGGAGGGGAAAGCACCAACCATGGCAGAGTCCA 650

Qy     601 GAAAGGCCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCGAATGAAAAT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     651 GAAAGGCCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCGAATGAAAAT 710

Qy     661 GAGCCACACCTTCCACATGGTGTTAGCCAGAGAAGAGGGGTCAGGGACTCTTCCCTGGTC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     711 GAGCCACACCTTCCACATGGTGTTAGCCAGAGAAGAGGGGTCAGGGACTCTTCCCTGGTC 770

Qy     721 GTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGGAAACTTCTTGGTGCTCAAAGCAAAG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     771 GTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGGAAACTTCTTGGTGCTCAAAGCAAAG 830

Qy     781 GAGATGGGCCTCCCAGTTGGGACAGCTGCCATCGCTCCCATCATTGCTGCTGTCAAGGAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     831 GAGATGGGCCTCCCAGTTGGGACAGCTGCCATCGCTCCCATCATTGCTGCTGTCAAGGAC 890

Qy     841 GGGAAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCCTCCA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     891 GGGAAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCCTCCA 950

Qy     901 GATCCTGGTGCTGCTTTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCATTCAACCCATC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     951 GATCCTGGTGCTGCTTTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCATTCAACCCATC 1010

Qy     961 TGTGAGAATGCCACCTTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCGTGGCCTTGGTG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1011 TGTGAGAATGCCACCTTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCGTGGCCTTGGTG 1070

Qy    1021 GTTCACATGGCCCCAGCATCTGTGCTTGTGGACAGCAGGTACCAGCAGTGGATGGAGAGG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1071 GTTCACATGGCCCCAGCATCTGTGCTTGTGGACAGCAGGTACCAGCAGTGGATGGAGAGG 1130

Qy    1081 TTTGGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1131 TTTGGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTT 1190

Qy    1141 CGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCTC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1191 CGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCTC 1250

Qy    1201 ACCAGTTTCCGCTGTAAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGGTTCAGGGTGAA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1251 ACCAGTTTCCGCTGTAAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGGTTCAGGGTGAA 1310

Qy    1261 TGCCTCCTCAAGTACCAGCTCCGTCCCAGGAGGGAGTGGCAGAGGGATGCCATTATTACT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1311 TGCCTCCTCAAGTACCAGCTCCGTCCCAGGAGGGAGTGGCAGAGGGATGCCATTATTACT 1370

Qy    1321 TGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGCGTG 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1371 TGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGCGTG 1430

Qy    1381 CAGGAGTACAGGAGGAGTGCGCAGGACGGCCCAGCCCCAGCAGAGAAAAGAAGTCAGTAC 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1431 CAGGAGTACAGGAGGAGTGCGCAGGACGGCCCAGCCCCAGCAGAGAAAAGAAGTCAGTAC 1490

Qy    1441 CCAGAAATCATCTTCCTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGT 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1491 CCAGAAATCATCTTCCTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGT 1550

Qy    1501 GCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACA 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1551 GCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACA 1610
```

```
Qy   1561 TTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1611 TTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCT 1670

Qy   1621 GCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGGGCTTGCCAAGTATCTTGCTG 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1671 GCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGGGCTTGCCAAGTATCTTGCTG 1730

Qy   1681 CAGAGAGAACGCGCCTTGGCATCTTTGGGAAAGCCGCTTCACCCTTTGCTGGTGGTTGCC 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1731 CAGAGAGAACGCGCCTTGGCATCTTTGGGAAAGCCGCTTCACCCTTTGCTGGTGGTTGCC 1790

Qy   1741 CCCAACCAGCTCAAAGCCTGGCTCCAGCAGTACCACAACCAGTGCCAGGAGGTCCTGCAC 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1791 CCCAACCAGCTCAAAGCCTGGCTCCAGCAGTACCACAACCAGTGCCAGGAGGTCCTGCAC 1850

Qy   1801 CACATCAGTATGATTCCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCTCCAGTCCTGCA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1851 CACATCAGTATGATTCCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCTCCAGTCCTGCA 1910

Qy   1861 GTGGAAAGATTGATCAGTTCGCTGTTGCGAACATGTGATTTGGAAGAGTTTCAGACCTGT 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1911 GTGGAAAGATTGATCAGTTCGCTGTTGCGAACATGTGATTTGGAAGAGTTTCAGACCTGT 1970

Qy   1921 CTGGTGCGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAA 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1971 CTGGTGCGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAA 2030

Qy   1981 GTGGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCTGGTCCGGATGGGGAAAGATGCC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2031 GTGGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCTGGTCCGGATGGGGAAAGATGCC 2090

Qy   2041 ACCCTCCTGATACATGAAGCCACCCTGGAAGATGGTTTGGAAGAGGAAGCAGTGGAAAAG 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2091 ACCCTCCTGATACATGAAGCCACCCTGGAAGATGGTTTGGAAGAGGAAGCAGTGGAAAAG 2150

Qy   2101 ACACACAGCACAACGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTCATT 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2151 ACACACAGCACAACGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTCATT 2210

Qy   2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCCCTCTTCAGCCCCAACTTCAGC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2211 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCCCTCTTCAGCCCCAACTTCAGC 2270

Qy   2221 GAGAAAGTGGGAGTTGCCTTTGACCACATGAAGGTCTGCTTTGGAGACTTTCCAACAATG 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2271 GAGAAAGTGGGAGTTGCCTTTGACCACATGAAGGTCTGCTTTGGAGACTTTCCAACAATG 2330

Qy   2281 CCCAAGCTGATTCCCCCACTGAAAGCCCTGTTTGCTGGCGACATCGAGGAGATGGAGGAG 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2331 CCCAAGCTGATTCCCCCACTGAAAGCCCTGTTTGCTGGCGACATCGAGGAGATGGAGGAG 2390

Qy   2341 CGCAGGGAGAAGCGGGAGCTGCGGCAGGTGCGGGCGGCCCTCCTGTCCAGGGAGCTGGCA 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2391 CGCAGGGAGAAGCGGGAGCTGCGGCAGGTGCGGGCGGCCCTCCTGTCCAGGGAGCTGGCA 2450

Qy   2401 GGCGGCCTGGAGGATGGGGAGCCTCAGCAGAAGCGGGCCCACACAGAGGAGCCACAGGCC 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2451 GGCGGCCTGGAGGATGGGGAGCCTCAGCAGAAGCGGGCCCACACAGAGGAGCCACAGGCC 2510

Qy   2461 AAGAAGGTCAGAGCCCAGTGA 2481
          |||||||||||||||||||||
Db   2511 AAGAAGGTCAGAGCCCAGTGA 2531
```

RESULT 4
ABN59829
ID ABN59829 standard; cDNA; 2992 BP.
XX
AC ABN59829;
XX
DT 28-JUN-2002 (first entry)
XX
DE Novel human coding sequence SEQ ID NO: 240.
XX
KW Human; antianaemic; vulnerary; antiinflammatory; immunomodulator;
KW antiinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
KW neuroprotective; antiparkinsonian; protein therapy; EST;
KW expressed sequence tag; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200222660-A2.
XX
PD 21-MAR-2002.
XX
PF 10-SEP-2001; 2001WO-US26015.
XX
PR 11-SEP-2000; 2000US-0659671.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
DR WPI; 2002-292408/33.
DR P-PSDB; ABB97416.
XX
PT An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis -
XX
PS Claim 1; SEQ ID NO 240; 509pp; English.
XX
CC The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a coding sequence of the
CC invention.
XX
SQ Sequence 2992 BP; 725 A; 807 C; 859 G; 601 T; 0 other;

```
  Query Match            100.0%;  Score 2481;  DB 24;  Length 2992;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 2481;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ATGTGGGCGCTTTGCTCGCTGCTGCGGTCCGCGGCCGGACGCACCATGTCGCAGGGACGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     75 ATGTGGGCGCTTTGCTCGCTGCTGCGGTCCGCGGCCGGACGCACCATGTCGCAGGGACGC 134

Qy     61 ACCATATCGCAGGCACCCGCCCGCCGCGAGCGGCCGCGCAAGGACCCGCTGCGGCACCTG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    135 ACCATATCGCAGGCACCCGCCCGCCGCGAGCGGCCGCGCAAGGACCCGCTGCGGCACCTG 194

Qy    121 CGCACGCGAGAGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAAACACCGTGTACCTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    195 CGCACGCGAGAGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAAACACCGTGTACCTG 254

Qy    181 CAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    255 CAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTC 314

Qy    241 AACCGGTATCTCTTCAACTGTGGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    315 AACCGGTATCTCTTCAACTGTGGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTA 374

Qy    301 AAGGTTGCTCGCCTGGACAACATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    375 AAGGTTGCTCGCCTGGACAACATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGC 434

Qy    361 TTAAGTGGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 420
```

```
Db    435 TTAAGTGGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 494

Qy    421 CCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGA 480

Db    495 CCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGA 554

Qy    481 ATAGAACTGGCTGTGCGGCCCCACTCTGCCCCAGAATACGAGGATGAAACCATGACAGTT 540

Db    555 ATAGAACTGGCTGTGCGGCCCCACTCTGCCCCAGAATACGAGGATGAAACCATGACAGTT 614

Qy    541 TACCAGATCCCCATACACAGTGAACAGAGGAGGGGAAAGCACCAACCATGGCAGAGTCCA 600

Db    615 TACCAGATCCCCATACACAGTGAACAGAGGAGGGGAAAGCACCAACCATGGCAGAGTCCA 674

Qy    601 GAAAGGCCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCGAATGAAAAT 660

Db    675 GAAAGGCCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCGAATGAAAAT 734

Qy    661 GAGCCACACCTTCCACATGGTGTTAGCCAGAGAAGAGGGGTCAGGGACTCTTCCCTGGTC 720

Db    735 GAGCCACACCTTCCACATGGTGTTAGCCAGAGAAGAGGGGTCAGGGACTCTTCCCTGGTC 794

Qy    721 GTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGGAAACTTCTTGGTGCTCAAAGCAAAG 780

Db    795 GTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGGAAACTTCTTGGTGCTCAAAGCAAAG 854

Qy    781 GAGATGGGCCTCCCAGTTGGGACAGCTGCCATCGCTCCCATCATTGCTGCTGTCAAGGAC 840

Db    855 GAGATGGGCCTCCCAGTTGGGACAGCTGCCATCGCTCCCATCATTGCTGCTGTCAAGGAC 914

Qy    841 GGGAAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCCTCCA 900

Db    915 GGGAAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCCTCCA 974

Qy    901 GATCCTGGTGCTGCTTTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCATTCAACCCATC 960

Db    975 GATCCTGGTGCTGCTTTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCATTCAACCCATC 1034

Qy    961 TGTGAGAATGCCACCTTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCGTGGCCTTGGTG 1020

Db   1035 TGTGAGAATGCCACCTTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCGTGGCCTTGGTG 1094

Qy   1021 GTTCACATGGCCCCAGCATCTGTGCTTGTGGACAGCAGGTACCAGCAGTGGATGGAGAGG 1080

Db   1095 GTTCACATGGCCCCAGCATCTGTGCTTGTGGACAGCAGGTACCAGCAGTGGATGGAGAGG 1154

Qy   1081 TTTGGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTT 1140

Db   1155 TTTGGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTT 1214

Qy   1141 CGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCTC 1200

Db   1215 CGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCTC 1274

Qy   1201 ACCAGTTTCCGCTGTAAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGGTTCAGGGTGAA 1260

Db   1275 ACCAGTTTCCGCTGTAAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGGTTCAGGGTGAA 1334

Qy   1261 TGCCTCCTCAAGTACCAGCTCCGTCCCAGGAGGGAGTGGCAGAGGGATGCCATTATTACT 1320

Db   1335 TGCCTCCTCAAGTACCAGCTCCGTCCCAGGAGGGAGTGGCAGAGGGATGCCATTATTACT 1394

Qy   1321 TGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGCGTG 1380

Db   1395 TGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGCGTG 1454

Qy   1381 CAGGAGTACAGGAGGAGTGCGCAGGACGGCCCAGCCCCAGCAGAGAAAAGAAGTCAGTAC 1440

Db   1455 CAGGAGTACAGGAGGAGTGCGCAGGACGGCCCAGCCCCAGCAGAGAAAAGAAGTCAGTAC 1514

Qy   1441 CCAGAAATCATCTTCCTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGT 1500

Db   1515 CCAGAAATCATCTTCCTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGT 1574

Qy   1501 GCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACA 1560

Db   1575 GCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACA 1634

Qy   1561 TTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCT 1620

Db   1635 TTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCT 1694

Qy   1621 GCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGGGCTTGCCAAGTATCTTGCTG 1680

Db   1695 GCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGGGCTTGCCAAGTATCTTGCTG 1754

Qy   1681 CAGAGAGAACGCGCCTTGGCATCTTTGGGAAAGCCGCTTCACCCTTTGCTGGTGGTTGCC 1740

Db   1755 CAGAGAGAACGCGCCTTGGCATCTTTGGGAAAGCCGCTTCACCCTTTGCTGGTGGTTGCC 1814

Qy   1741 CCCAACCAGCTCAAAGCCTGGCTCCAGCAGTACCACAACCAGTGCCAGGAGGTCCTGCAC 1800

Db   1815 CCCAACCAGCTCAAAGCCTGGCTCCAGCAGTACCACAACCAGTGCCAGGAGGTCCTGCAC 1874

Qy   1801 CACATCAGTATGATTCCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCTCCAGTCCTGCA 1860

Db   1875 CACATCAGTATGATTCCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCTCCAGTCCTGCA 1934

Qy   1861 GTGGAAAGATTGATCAGTTCGCTGTTGCGAACATGTGATTTGGAAGAGTTTCAGACCTGT 1920

Db   1935 GTGGAAAGATTGATCAGTTCGCTGTTGCGAACATGTGATTTGGAAGAGTTTCAGACCTGT 1994

Qy   1921 CTGGTGCGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAA 1980

Db   1995 CTGGTGCGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAA 2054

Qy   1981 GTGGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCTGGTCCGGATGGGGAAAGATGCC 2040

Db   2055 GTGGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCTGGTCCGGATGGGGAAAGATGCC 2114

Qy   2041 ACCCTCCTGATACATGAAGCCACCCTGGAAGATGGTTTGGAAGAGGAAGCAGTGGAAAAG 2100

Db   2115 ACCCTCCTGATACATGAAGCCACCCTGGAAGATGGTTTGGAAGAGGAAGCAGTGGAAAAG 2174

Qy   2101 ACACACAGCACAACGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTCATT 2160

Db   2175 ACACACAGCACAACGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTCATT 2234

Qy   2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCCCTCTTCAGCCCCAACTTCAGC 2220

Db   2235 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCCCTCTTCAGCCCCAACTTCAGC 2294

Qy   2221 GAGAAAGTGGGAGTTGCCTTTGACCACATGAAGGTCTGCTTTGGAGACTTTCCAACAATG 2280

Db   2295 GAGAAAGTGGGAGTTGCCTTTGACCACATGAAGGTCTGCTTTGGAGACTTTCCAACAATG 2354

Qy   2281 CCCAAGCTGATTCCCCCACTGAAAGCCCTGTTTGCTGGCGACATCGAGGAGATGGAGGAG 2340

Db   2355 CCCAAGCTGATTCCCCCACTGAAAGCCCTGTTTGCTGGCGACATCGAGGAGATGGAGGAG 2414

Qy   2341 CGCAGGGAGAAGCGGGAGCTGCGGCAGGTGCGGGCGGCCCTCCTGTCCAGGGAGCTGGCA 2400

Db   2415 CGCAGGGAGAAGCGGGAGCTGCGGCAGGTGCGGGCGGCCCTCCTGTCCAGGGAGCTGGCA 2474

Qy   2401 GGCGGCCTGGAGGATGGGGAGCCTCAGCAGAAGCGGGCCCACACAGAGGAGCCACAGGCC 2460

Db   2475 GGCGGCCTGGAGGATGGGGAGCCTCAGCAGAAGCGGGCCCACACAGAGGAGCCACAGGCC 2534

Qy   2461 AAGAAGGTCAGAGCCCAGTGA 2481

Db   2535 AAGAAGGTCAGAGCCCAGTGA 2555

RESULT 5
AAA52810
```

```
ID   AAA52810 standard; cDNA; 2478 BP.
XX
AC   AAA52810;
XX
DT   20-SEP-2000  (first entry)
XX
DE   Human sulphatase G cDNA.
XX
KW   Human sulphatase G; hSG; chromosome 17p11.2; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..2478
FT                   /*tag= a
FT                   /partial
FT                   /product= "hSG"
XX
PN   WO200034327-A1.
XX
PD   15-JUN-2000.
XX
PF   09-DEC-1999;   99WO-AU01092.
XX
PR   09-DEC-1998;   98AU-0007624.
XX
PA   (WOME-) WOMEN'S & CHILDREN'S HOSPITAL.
XX
PI   Hopwood JJ,  Litjens T,  Hu RL;
XX
DR   WPI; 2000-431273/37.
DR   P-PSDB; AAY99850.
XX
PT   Novel isolated DNA sequence which encodes human sulfatase G or its
PT   fragment useful in gene therapy for treating patients suffering from
PT   sulfatase deficiency  -
XX
PS   Claim 2; Page 29-30; 33pp; English.
XX
CC   The present sequence encodes human sulphatase G (hSG). hSG is
CC   not a member of the well-characterised CTPSR sulphatase family.
CC   It belongs to a family showing sequence similarity to a sulphatase
CC   from the marine bacterium Alteromonas carrageenovora. The hSG gene
CC   contains 23 exons and is located at chromosome 17p11.2. The present
CC   sequence is clone lambda29.1 from a human testes cDNA library. It was
CC   isolated using human EST sequences with sequence similarity to the
CC   non-CTPSR family as a probe to screen the library. The cDNA insert was
CC   subcloned and the DNA sequence of both strands was determined. The
CC   sequence may be used to treat a patient suffering from hSG deficiency
CC   by replacing, repairing, or compensating for a DNA sequence within that
CC   patient's genome.
XX
SQ   Sequence 2478 BP; 587 A; 686 C; 709 G; 496 T; 0 other;

  Query Match             99.6%;  Score 2471.6;  DB 21;  Length 2478;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 2474;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 ATGTGGGCGCTTTGCTCGCTGCTGCGGTCCGCGGCCGGACGCACCATGTCGCAGGGACGC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGTGGGCGCTTTGCTCGCTGCTGCGGTCCGCGGCCGGACGCACCATGTCGCAGGGACGC 60

Qy      61 ACCATATCGCAGGCACCCGCCCGCCGCGAGCGGCCGCGCAAGGACCCGCTGCGGCACCTG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACCATATCGCAGGCACCCGCCCGCCGCGAGCGGCCGCGCAAGGACCCGCTGCGGCACCTG 120

Qy     121 CGCACGCGAGAGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAAACACCGTGTACCTG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CGCACGCGAGAGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAAACACCGTGTACCTG 180

Qy     181 CAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTC 240

Qy     241 AACCGGTATCTCTTCAACTGTGGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AACCGGTATCTCTTCAACTGTGGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTA 300

Qy     301 AAGGTTGCTCGCCTGGACAACATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 AAGGTTGCTCGCCTGGACAACATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGC 360

Qy     361 TTAAGTGGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTAAGTGGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 420

Qy     421 CCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGA 480

Qy     481 ATAGAACTGGCTGTGCGGCCCCACTCTGCCCCAGAATACGAGGATGAAACCATGACAGTT 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 ATAGAACTGGCTGTGCGGCCCCACTCTGCCCCAGAATACGAGGATGAAACCATGACAGTT 540

Qy     541 TACCAGATCCCCATACACAGTGAACAGAGGAGGGGAAAGCACCAACCATGGCAGAGTCCA 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 TACCAGATCCCCATACACAGTGAACAGAGGAGGGGAAAGCACCAACCATGGCAGAGTCCA 600

Qy     601 GAAAGGCCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCGAATGAAAAT 660
           |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     601 GAAAGGCCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTTGAATGAAAAT 660

Qy     661 GAGCCACACCTTCCACATGGTGTTAGCCAGAGAAGAGGGGTCAGGGACTCTTCCCTGGTC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GAGCCACACCTTCCACATGGTGTTAGCCAGAGAAGAGGGGTCAGGGACTCTTCCCTGGTC 720

Qy     721 GTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGGAAACTTCTTGGTGCTCAAAGCAAAG 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGGAAACTTCTTGGTGCTCAAAGCAAAG 780

Qy     781 GAGATGGGCCTCCCAGTTGGGACAGCTGCCATCGCTCCCATCATTGCTGCTGTCAAGGAC 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GAGATGGGCCTCCCAGTTGGGACAGCTGCCATCGCTCCCATCATTGCTGCTGTCAAGGAC 840

Qy     841 GGGAAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCCTCCA 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 GGGAAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCCTCCA 900

Qy     901 GATCCTGGTGCTGCTTTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCATTCAACCCATC 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 GATCCTGGTGCTGCTTTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCATTCAACCCATC 960

Qy     961 TGTGAGAATGCCACCTTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCGTGGCCTTGGTG 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 TGTGAGAATGCCACCTTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCGTGGCCTTGGTG 1020

Qy    1021 GTTCACATGGCCCCAGCATCTGTGCTTGTGGACAGCAGGTACCAGCAGTGGATGGAGAGG 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GTTCACATGGCCCCAGCATCTGTGCTTGTGGACAGCAGGTACCAGCAGTGGATGGAGAGG 1080

Qy    1081 TTTGGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 TTTGGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTT 1140

Qy    1141 CGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCTC 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 CGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCTC 1200

Qy    1201 ACCAGTTTCCGCTGTAAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGGTTCAGGGTGAA 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 ACCAGTTTCCGCTGTAAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGGTTCAGGGTGAA 1260

Qy    1261 TGCCTCCTCAAGTACCAGCTCCGTCCCAGGAGGGAGTGGCAGAGGGATGCCATTATTACT 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 TGCCTCCTCAAGTACCAGCTCCGTCCCAGGAGGGAGTGGCAGAGGGATGCCATTATTACT 1320
```

```
Qy   1321 TGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGCGTG 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 TGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGCGTG 1380

Qy   1381 CAGGAGTACAGGAGGAGTGCGCAGGACGGCCCAGCCCCAGCAGAGAAAAGAAGTCAGTAC 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CAGGAGTACAGGAGGAGTGCGCAGGACGGCCCAGCCCCAGCAGAGAAAAGAAGTCAGTAC 1440

Qy   1441 CCAGAAATCATCTTCCTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGT 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 CCAGAAATCATCTTCCTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGT 1500

Qy   1501 GCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACA 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACG 1560

Qy   1561 TTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 TTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCT 1620

Qy   1621 GCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGGGCTTGCCAAGTATCTTGCTG 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 ACTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGGGCTTGCCAAGTATCTTGCTG 1680

Qy   1681 CAGAGAGAACGCGCCTTGGCATCTTTGGGAAAGCCGCTTCACCCTTTGCTGGTGGTTGCC 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 CAGAGAGAACGCGCCTTGGCATCTTTGGGAAAGCCGCTTCACCCTTTGCTGGTGGTTGCC 1740

Qy   1741 CCCAACCAGCTCAAAGCCTGGCTCCAGCAGTACCACAACCAGTGCCAGGAGGTCCTGCAC 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 CCCAACCAGCTCAAAGCCTGGCTCCAGCAGTACCACAACCAGTGCCAGGAGGTCCTGCAC 1800

Qy   1801 CACATCAGTATGATTCCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCTCCAGTCCTGCA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 CACATCAGTATGATTCCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCTCCAGTCCTGCA 1860

Qy   1861 GTGGAAAGATTGATCAGTTCGCTGTTGCGAACATGTGATTTGGAAGAGTTTCAGACCTGT 1920
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   1861 GTGGAAAGATTGATCAGTTCCCTGTTGCGAACATGTGATTTGGAAGAGTTTCAGACCTGT 1920

Qy   1921 CTGGTGCGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAA 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CTGGTGCGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAA 1980

Qy   1981 GTGGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCTGGTCCGGATGGGGAAAGATGCC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 GTGGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCTGGTCCGGATGGGGAAAGATGCC 2040

Qy   2041 ACCCTCCTGATACATGAAGCCACCCTGGAAGATGGTTTGGAAGAGGAAGCAGTGGAAAAG 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 ACCCTCCTGATACATGAAGCCACCCTGGAAGATGGTTTGGAAGAGGAAGCAGTGGAAAAG 2100

Qy   2101 ACACACAGCACAACGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTCATT 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 ACACACAGCACAACGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTCATT 2160

Qy   2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCCCTCTTCAGCCCCAACTTCAGC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCCCTCTTCAGCCCCAACTTCAGC 2220

Qy   2221 GAGAAAGTGGGAGTTGCCTTTGACCACATGAAGGTCTGCTTTGGAGACTTTCCAACAATG 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 GAGAAAGTGGGAGTTGCCTTTGACCACATGAAGGTCTGCTTTGGAGACTTTCCAACAATG 2280

Qy   2281 CCCAAGCTGATTCCCCCACTGAAAGCCCTGTTTGCTGGCGACATCGAGGAGATGGAGGAG 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 CCCAAGCTGATTCCCCCACTGAAAGCCCTGTTTGCTGGCGACATCGAGGAGATGGAGGAG 2340

Qy   2341 CGCAGGGAGAAGCGGGAGCTGCGGCAGGTGCGGGCGGCCCTCCTGTCCAGGGAGCTGGCA 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 CGCAGGGAGAAGCGGGAGCTGCGGCAGGTGCGGGCGGCCCTCCTGTCCAGGGAGCTGGCA 2400

Qy   2401 GGCGGCCTGGAGGATGGGGAGCCTCAGCAGAAGCGGGCCCACACAGAGGAGCCACAGGCC 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 GGCGGCCTGGAGGATGGGGAGCCTCAGCAGAAGCGGGCCCACACAGAGGAGCCACAGGCC 2460

Qy   2461 AAGAAGGTCAGAGCCCAG 2478
          ||||||||||||||||||
Db   2461 AAGAAGGTCAGAGCCCAG 2478
```

RESULT 6
AAS99132
ID AAS99132 standard; cDNA; 2908 BP.
XX
AC AAS99132;
XX
DT 12-MAR-2002 (first entry)
XX
DE Chimpanzee ELAC2 cDNA.
XX
KW Human; mouse; HPC2; prostate cancer; neoplastic growth; cytostatic; ss;
KW gene therapy; prostate cancer predisposing gene; chimpanzee; gorilla;
KW sequencing primer; PCR primer.
XX
OS Pan troglodytes.
XX
PN WO200185911-A2.
XX
PD 15-NOV-2001.
XX
PF 07-MAY-2001; 2001WO-US14602.
XX
PR 05-MAY-2000; 2000US-0564805.
XX
PA (MYRI-) MYRIAD GENETICS INC.
PA (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI Tavtigian SV, Teng DHF, Simard J, Rommens JM;
XX
DR WPI; 2002-066599/09.
DR P-PSDB; AAU73592.
XX
PT Novel nucleic acid sequence encoding HPC2 polypeptide, which is marker
PT for prostate cancer, is useful in gene therapy techniques to restore
PT HPC2 normal levels by which neoplastic growth is suppressed in
PT recipient cell -
XX
PS Claim 87; Page 198-201; 239pp; English.
XX
CC The invention relates to a human prostate cancer predisposing gene coding
CC for an HPC2 polypeptide. The DNA and protein sequences are useful as
CC diagnostic reagents for identifying a mutant HPC2 nucleotide sequence in
CC a suspected mutant HPC2 allele by comparing the sequence of the suspected
CC mutant HPC2 allele with a wild-type HPC2 sequence. The sequences are also
CC useful for detecting an alteration in HPC2, where the alteration is
CC associated with cancer in a human. The method involves analysing an HPC2
CC gene or an HPC2 gene expression product from a tissue of the human. The
CC HPC2 gene is useful as a marker for prostate cancer and can be used in
CC gene therapy techniques to suppress neoplastic growth of recipient cells
CC which carry the mutant HPC2 allele. The sequences represent primers used
CC in the methods of the invention, cDNA encoding human and mouse HPC2 and
CC cDNA encoding HPC2 paralogues and orthologues.
XX
SQ Sequence 2908 BP; 712 A; 788 C; 819 G; 589 T; 0 other;

Query Match 99.0%; Score 2455.4; DB 24; Length 2908;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 2465; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
Qy      1 ATGTGGGCGCTTTGCTCGCTGCTGCGGTCCGCGGCCGGACGCACCATGTCGCAGGGACGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGTGGGCGCTTTGCTCGCTGCTGCGGTCCGCGGCCGGACGCACCATGTCGCAGGGACGC 60

Qy     61 ACCATATCGCAGGCACCCGCCCGCCGCGAGCGGCCGCGCAAGGACCCGCTGCGGCACCTG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     61 ACCATATCGCAGGCACCCGCCCGCCGCGAGCGGCCGCGCAAGGACCCGCTGCGGCACCTG 120

Qy    121 CGCACGCGAGAGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAAACACCGTGTACCTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CGCACGCGAGAGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAAACACCGTGTACCTG 180

Qy    181 CAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTC 240

Qy    241 AACCGGTATCTCTTCAACTGTGGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTA 300
          |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    241 AACCGGTATCTCTTCAACTGTGGAGAAGGCATTCAGAGACTCATGCAGGAGCACAAGTTA 300

Qy    301 AAGGTTGCTCGCCTGGACAACATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AAGGTTGCTCGCCTGGACAACATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGC 360

Qy    361 TTAAGTGGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTAAGTGGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 420

Qy    421 CCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGA 480

Qy    481 ATAGAACTGGCTGTGCGGCCCCACTCTGCCCCAGAATACGAGGATGAAACCATGACAGTT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ATAGAACTGGCTGTGCGGCCCCACTCTGCCCCAGAATACGAGGATGAAACCATGACAGTT 540

Qy    541 TACCAGATCCCCATACACAGTGAACAGAGGAGGGGAAAGCACCAACCATGGCAGAGTCCA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TACCAGATCCCCATACACAGTGAACAGAGGAGGGGAAAGCACCAACCATGGCAGAGTCCA 600

Qy    601 GAAAGGCCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCGAATGAAAAT 660
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    601 GAAAGGCCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCAAATGAAAAT 660

Qy    661 GAGCCACACCTTCCACATGGTGTTAGCCAGAGAAGAGGGGTCAGGGACTCTTCCCTGGTC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GAGCCACACCTTCCACATGGTGTTAGCCAGAGAAGAGGGGTCAGGGACTCTTCCCTGGTC 720

Qy    721 GTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGGAAACTTCTTGGTGCTCAAAGCAAAG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGGAAACTTCTTGGTGCTCAAAGCAAAG 780

Qy    781 GAGATGGGCCTCCCAGTTGGGACAGCTGCCATCGCTCCCATCATTGCTGCTGTCAAGGAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GAGATGGGCCTCCCAGTTGGGACAGCTGCCATCGCTCCCATCATTGCTGCTGTCAAGGAC 840

Qy    841 GGGAAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCCTCCA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 GGGAAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCCTCCA 900

Qy    901 GATCCTGGTGCTGCTTTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCATTCAACCCATC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GATCCTGGTGCTGCTTTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCATTCAACCCATC 960

Qy    961 TGTGAGAATGCCACCTTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCGTGGCCTTGGTG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TGTGAGAATGCCACCTTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCGTGGCCTTGGTG 1020

Qy   1021 GTTCACATGGCCCCAGCATCTGTGCTTGTGGACAGCAGGTACCAGCAGTGGATGGAGAGG 1080
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GTTCACATGGCCCCAGAATCTGTGCTTGTGGACAGCAGGTACCAGCAGTGGATGGAGAGG 1080

Qy   1081 TTTGGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TTTGGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTT 1140

Qy   1141 CGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCTC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCTC 1200

Qy   1201 ACCAGTTTCCGCTGTAAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGGTTCAGGGTGAA 1260
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ACCAGTTTCCCCTGTAAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGGTTCAGGGTGAA 1260

Qy   1261 TGCCTCCTCAAGTACCAGCTCCGTCCCAGGAGGGAGTGGCAGAGGGATGCCATTATTACT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 TGCCTCCTCAAGTACCAGCTCCGTCCCAGGAGGGAGTGGCAGAGGGATGCCATTATTACT 1320

Qy   1321 TGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGCGTG 1380
          |||||||||||||||||| |||||||||||||||||||||||||||||||||||||| |||
Db   1321 TGCAATCCTGAGGAATTCATAATTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGTGTG 1380

Qy   1381 CAGGAGTACAGGAGGAGTGCGCAGGACGGCCCAGCCCCAGCAGAGAAAAGAAGTCAGTAC 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CAGGAGTACAGGAGGAGTGCGCAGGACGGCCCAGCCCCAGCAGAGAAAAGAAGTCAGTAC 1440

Qy   1441 CCAGAAATCATCTTCCTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGT 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 CCAGAAATCATCTTCCTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGT 1500

Qy   1501 GCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACA 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACG 1560

Qy   1561 TTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 TTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCT 1620

Qy   1621 GCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGGGCTTGCCAAGTATCTTGCTG 1680
          |||||||||||||||||||||||||||||||||||||||||||||||| || ||||||||||
Db   1621 GCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGGGCTTGCTAAATATCTTGCTG 1680

Qy   1681 CAGAGAGAACGCGCCTTGGCATCTTTGGGAAAGCCGCTTCACCCTTTGCTGGTGGTTGCC 1740
          |||||||||| |||||||||||||||||||| |||||||||||||||||||||||||||||
Db   1681 CAGAGAGAACGAGCCTTGGCATCTTTGGGAAAGCCCTTTCACCCTTTGCTGGTGGTTGCC 1740

Qy   1741 CCCAACCAGCTCAAAGCCTGGCTCCAGCAGTACCACAACCAGTGCCAGGAGGTCCTGCAC 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 CCCAACCAGCTCAAAGCCTGGCTCCAGCAGTACCACAACCAGTGCCAGGAGGTCCTGCAC 1800

Qy   1801 CACATCAGTATGATTCCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCTCCAGTCCTGCA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 CACATCAGTATGATTCCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCTCCAGTCCTGCA 1860

Qy   1861 GTGGAAAGATTGATCAGTTCGCTGTTGCGAACATGTGATTTGGAAGAGTTTCAGACCTGT 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 GTGGAAAGATTGATCAGTTCGCTGTTGCGAACATGTGATTTGGAAGAGTTTCAGACCTGT 1920

Qy   1921 CTGGTGCGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAA 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CTGGTGCGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAA 1980

Qy   1981 GTGGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCTGGTCCGGATGGGGAAAGATGCC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 GTGGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCTGGTCCGGATGGGGAAAGATGCC 2040

Qy   2041 ACCCTCCTGATACATGAAGCCACCCTGGAAGATGGTTTGGAAGAGGAAGCAGTGGAAAAG 2100
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   2041 ACCCTCCTGATACATGAAGCCACCCTGGAAGACGGTTTGGAAGAGGAAGCAGTGGAAAAG 2100

Qy   2101 ACACACAGCACAACGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTCATT 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 ACACACAGCACAACGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTCATT 2160

Qy   2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCCCTCTTCAGCCCCAACTTCAGC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCCCTCTTCAGCCCCAACTTCAAC 2220

Qy   2221 GAGAAAGTGGGAGTTGCCTTTGACCACATGAAGGTCTGCTTTGGAGACTTTCCAACAATG 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 GAGAAAGTGGGAGTTGCCTTTGACCACATGAAGGTCTGCTTTGGAGACTTTGCAACAATG 2280
```

```
Qy    2281 CCCAAGCTGATTCCCCCACTGAAAGCCCTGTTTGCTGGCGACATCGAGGAGATGGAGGAG 2340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2281 CCCAAGCTGATTCCCCCACTGAAAGCCCTGTTTGCTGGCGACATCGAGGAGATGGAGGAG 2340

Qy    2341 CGCAGGGAGAAGCGGGAGCTGCGGCAGGTGCGGGCGGCCCTCCTGTCCAGGGAGCTGGCA 2400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2341 CGCAGGGAGAAGCGGGAGCTGCGGCAGGTGCGGGCGGCCCTCCTGTCCAGGGAGCTGGCA 2400

Qy    2401 GGCGGCCTGGAGGATGGGGAGCCTCAGCAGAAGCGGGCCCACACAGAGGAGCCACAGGCC 2460
           ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    2401 GGCGGCCTGGAGGATGGGGAGCCTCAGCAGAAACGGGCCCACACAGAGGAGCCACAGGCC 2460

Qy    2461 AAGAAGGTCAGAGCCCAGTGA 2481
           |||||||||||||||||||||
Db    2461 AAGAAGGTCAGAGCCCAGTGA 2481
```

```
RESULT 7
AAH14250
ID   AAH14250 standard; cDNA; 2976 BP.
XX
AC   AAH14250;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA sequence SEQ ID NO:11557.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;   99JP-0248036.
PR   27-AUG-1999;   99JP-0300253.
PR   11-JAN-2000; 2000JP-0118776.
PR   02-MAY-2000; 2000JP-0183767.
PR   09-JUN-2000; 2000JP-0241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602
PT   full-length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full-length cDNAs -
XX
PS   Claim 8; SEQ ID 11557; 2537pp + CD ROM; English.
XX
CC   The present invention describes primer sets for synthesising 5602
CC   full-length cDNAs defined in the specification. Where a primer set
CC   comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC   to the complementary strand of a polynucleotide which comprises one of
CC   the 5602 nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC   AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 2976 BP; 712 A; 807 C; 856 G; 601 T; 0 other;
```

```
  Query Match             98.7%;  Score 2448.2;  DB 22;  Length 2976;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 2474;  Conservative    0;  Mismatches    3;  Indels    4;  Gaps    2;

Qy       1 ATGTGGGCGCTTTGCTCGCTGCTGCGGTCCGCGGCCGGACGCACCATGTCGCAGGGACGC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75 ATGTGGGCGCTTTGCTCGCTGCTGCGGTCCGCGGCCGGACGCACCATGTCGCAGGGACGC 134

Qy      61 ACCATATCGCAGGCACCCGCCCGCCGCGAGCGGCCGCGCAAGGACCCGCTGCGGCACCTG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     135 ACCATATCGCAGGCACCCGCCCGCCGCGAGCGGCCGCGCAAGGACCCGCTGCGGCACCTG 194

Qy     121 CGCACGCGAGAGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAAACACCGTGTACCTG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     195 CGCACGCGAGAGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAAACACCGTGTACCTG 254

Qy     181 CAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     255 CAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTC 314

Qy     241 AACCGGTATCTCTTCAACTGTGGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     315 AACCGGTATCTCTTCAACTGTGGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTA 374

Qy     301 AAGGTTGCTCGCCTGGACAACATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     375 AAGGTTGCTCGCCTGGACAACATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGC 434

Qy     361 TTAAGTGGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     435 TTAAGTGGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 494

Qy     421 CCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     495 CCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGA 554

Qy     481 ATAGAACTGGCTGTGCGGCCCCACTCTGCCCCAGAATACGAGGATGAAACCATGACAGTT 540
           |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db     555 ATAGAACTGGCTATGCGGCCCCACTCTGCCCCAGAATACGAGGATGAAACCATGACAGTT 614

Qy     541 TACCAGATCCCCATACACAGTGAACAGAGGAGGGGAAAGCACCAACCATGGCAGAGTCCA 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     615 TACCAGATCCCCATACACAGTGAACAGAGGAGGGGAAAGCACCAACCATGGCAGAGTCCA 674

Qy     601 GAAAGGCCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCGAATGAAAAT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     675 GAAAGGCCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCGAATGAAAAT 734

Qy     661 GAGCCACACCTTCCACATGGTGTTAGCCAGAGAAGAGGGGTCAGGGACTCTTCCCTGGTC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     735 GAGCCACACCTTCCACATGGTGTTAGCCAGAGAAGAGGGGTCAGGGACTCTTCCCTGGTC 794

Qy     721 GTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGGAAACTTCTTGGTGCTCAAAGCAAAG 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     795 GTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGGAAACTTCTTGGTGCTCAAAGCAAAG 854

Qy     781 GAGATGGGCCTCCCAGTTGGGACAGCTGCCATCGCTCCCATCATTGCTGCTGTCAAGGAC 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     855 GAGATGGGCCTCCCAGTTGGGACAGCTGCCATCGCTCCCATCATTGCTGCTGTCAAGGAC 914

Qy     841 GGGAAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCCTCCA 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     915 GGGAAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCCTCCA 974
```

```
Qy    901 GATCCTGGTGCTGCTTTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCATTCAACCCATC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    975 GATCCTGGTGCTGCTTTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCATTCAACCCATC 1034

Qy    961 TGTGAGAATGCCACCTTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCGTGGCCTTGGTG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1035 TGTGAGAATGCCACCTTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCGTGGCCTTGGTG 1094

Qy   1021 GTTCACATGGCCCCAGCATCTGTGCTTGTGGACAGCAGGTACCAGCAGTGGATGGAGAGG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1095 GTTCACATGGCCCCAGCATCTGTGCTTGTGGACAGCAGGTACCAGCAGTGGATGGAGAGG 1154

Qy   1081 TTTGGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1155 TTTGGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTT 1214

Qy   1141 CGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCTC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1215 CGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCTC 1274

Qy   1201 ACCAGTTTCCGCTGTAAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGGTTCAGGGTGAA 1260
          ||||||||||||||    ||||||||||||||||||||||||||||||||||||||||||
Db   1275 ACCAGTTTCCGCTGT---AAGGAGGGCCCCACCCTCAGTGTGCCCATGGTTCAGGGTGAA 1331

Qy   1261 TGCCTCCTCAAGTACCAGCTCCGTCCCAGGAGGGAGTGGCAGAGGGATGCCATTATTACT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1332 TGCCTCCTCAAGTACCAGCTCCGTCCCAGGAGGGAGTGGCAGAGGGATGCCATTATTACT 1391

Qy   1321 TGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGCGTG 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1392 TGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGCGTG 1451

Qy   1381 CAGGAGTACAGGAGGAGTGCGCAGGACGGCCCAGCCCCAGCAGAGAAAAGAAGTCAGTAC 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1452 CAGGAGTACAGGAGGAGTGCGCAGGACGGCCCAGCCCCAGCAGAGAAAAGAAGTCAGTAC 1511

Qy   1441 CCAGAAATCATCTTCCTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGT 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1512 CCAGAAATCATCTTCCTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGT 1571

Qy   1501 GCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACA 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1572 GCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACA 1631

Qy   1561 TTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1632 TTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCT 1691

Qy   1621 GCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGGGCTTGCCAAGTATCTTGCTG 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1692 GCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGGGCTTGCCAAGTATCTTGCTG 1751

Qy   1681 CAGAGAGAACGCGCCTTGGCATCTTTGGGAAAGCCGCTTCACCCTTTGCTGGTGGTTGCC 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1752 CAGAGAGAACGCGCCTTGGCATCTTTGGGAAAGCCGCTTCACCCTTTGCTGGTGGTTGCC 1811

Qy   1741 CCCAACCAGCTCAAAGCCTGGCTCCAGCAGTACCACAACCAGTGCCAGGAGGTCCTGCAC 1800
          ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db   1812 CCCAACCAGCTCAAAGCCTGGCTCCAGCAGTACTACAACCAGTGCCAGGAGGTCCTGCAC 1871

Qy   1801 CACATCAGTATGATTCCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCTCCAGTCCTGCA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1872 CACATCAGTATGATTCCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCTCCAGTCCTGCA 1931

Qy   1861 GTGGAAAGATTGATCAGTTCGCTGTTGCGAACATGTGATTTGGAAGAGTTTCAGACCTGT 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1932 GTGGAAAGATTGATCAGTTCGCTGTTGCGAACATGTGATTTGGAAGAGTTTCAGACCTGT 1991

Qy   1921 CTGGTGCGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAA 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1992 CTGGTGCGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAA 2051

Qy   1981 GTGGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCTGGTCCGGATGGGGAAAGATGCC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2052 GTGGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCTGGTCCGGATGGGGAAAGATGCC 2111

Qy   2041 ACCCTCCTGATACATGAAGCCACCCTGGAAGATGGTTTGGAAGAGGAAGCAGTGGAAAAG 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2112 ACCCTCCTGATACATGAAGCCACCCTGGAAGATGGTTTGGAAGAGGAAGCAGTGGAAAAG 2171

Qy   2101 ACACACAGCACAACGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTCATT 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2172 ACACACAGCACAACGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTCATT 2231

Qy   2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCCCTCTTCAGCCCCAACTTCAGC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2232 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCCCTCTTCAGCCCCAACTTCAGC 2291

Qy   2221 GAGAAAGTGGGAGTTGCCTTTGACCACATGAAGGTCTGCTTTGGAGACTTTCCAACAATG 2280
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   2292 GAGAAAGTGGGAGTTGCCTTTGACCACATGAAGGTCTGCCTTGGAGACTTTCCAACAATG 2351

Qy   2281 CCCAAGCTGATTCCCCCACTGAAAGCCCTGTTTGCTGGCGACATCGAGGAGATGGAGGAG 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2352 CCCAAGCTGATTCCCCCACTGAAAGCCCTGTTTGCTGGCGACATCGAGGAGATGGAGGAG 2411

Qy   2341 CGCAGGGAGAAGCGGGAGCTGCGGCAGGTGCGGGCGGCCCTCCTGTCCAGGGAGCTGGCA 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2412 CGCAGGGAGAAGCGGGAGCTGCGGCAGGTGCGGGCGGCCCTCCTGTCCAGGGAGCTGGCA 2471

Qy   2401 GGCGGCCTGGAGGATGGGGAGCCTCAGCAGAAGCGGGCCCACACAGAGGAGCCACAGGCC 2460
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   2472 GGCGGCCTGGAGGAT-GGGAGCCTCAGCAGAAGCGGGCCCACACAGAGGAGCCACAGGCC 2530

Qy   2461 AAGAAGGTCAGAGCCCAGTGA 2481
          |||||||||||||||||||||
Db   2531 AAGAAGGTCAGAGCCCAGTGA 2551
```

RESULT 8
AAS99133
ID AAS99133 standard; cDNA; 2892 BP.
XX
AC AAS99133;
XX
DT 12-MAR-2002 (first entry)
XX
DE Gorilla ELAC2 cDNA.
XX
KW Human; mouse; HPC2; prostate cancer; neoplastic growth; cytostatic; ss;
KW gene therapy; prostate cancer predisposing gene; chimpanzee; gorilla;
KW sequencing primer; PCR primer.
XX
OS Gorilla gorilla.
XX
PN WO200185911-A2.
XX
PD 15-NOV-2001.
XX
PF 07-MAY-2001; 2001WO-US14602.
XX
PR 05-MAY-2000; 2000US-0564805.
XX
PA (MYRI-) MYRIAD GENETICS INC.
PA (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI Tavtigian SV, Teng DHF, Simard J, Rommens JM;
XX
DR WPI; 2002-066599/09.
DR P-PSDB; AAU73593.
XX
PT Novel nucleic acid sequence encoding HPC2 polypeptide, which is marker
PT for prostate cancer, is useful in gene therapy techniques to restore
PT HPC2 normal levels by which neoplastic growth is suppressed in
PT recipient cell -
XX

PS Claim 92; Page 204-207; 239pp; English.
XX
CC The invention relates to a human prostate cancer predisposing gene coding
CC for an HPC2 polypeptide. The DNA and protein sequences are useful as
CC diagnostic reagents for identifying a mutant HPC2 nucleotide sequence in
CC a suspected mutant HPC2 allele by comparing the sequence of the suspected
CC mutant HPC2 allele with a wild-type HPC2 sequence. The sequences are also
CC useful for detecting an alteration in HPC2, where the alteration is
CC associated with cancer in a human. The method involves analysing an HPC2
CC gene or an HPC2 gene expression product from a tissue of the human. The
CC HPC2 gene is useful as a marker for prostate cancer and can be used in
CC gene therapy techniques to suppress neoplastic growth of recipient cells
CC which carry the mutant HPC2 allele. The sequences represent primers used
CC in the methods of the invention, cDNA encoding human and mouse HPC2 and
CC cDNA encoding HPC2 paralogues and orthologues.
XX
SQ Sequence 2892 BP; 704 A; 787 C; 815 G; 586 T; 0 other;

Query Match 98.5%; Score 2442.6; DB 24; Length 2892;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 2457; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

```
Qy       1 ATGTGGGCGCTTTGCTCGCTGCTGCGGTCCGCGGCCGGACGCACCATGTCGCAGGGACGC 60
Db       1 ATGTGGGCGCTTTGCTCGCTGCTGCGGTCCGCGGCCGGACGCACCATGTCGCAGGGACGC 60

Qy      61 ACCATATCGCAGGCACCCGCCCGCCGCGAGCGGCCGCGCAAGGACCCGCTGCGGCACCTG 120
Db      61 ACCATATCGCAGGCACCCGCCCGCCGCGAGCGGCCGCGCAAGGACCCGCTGCGGCACCTG 120

Qy     121 CGCACGCGAGAGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAAACACCGTGTACCTG 180
Db     121 CGCACGCGAGAGAAGCGCGGACCGTCGGGGTGCTCCGGGGGCCCAAACACCGTGTACCTG 180

Qy     181 CAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTC 240
Db     181 CAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTC 240

Qy     241 AACCGGTATCTCTTCAACTGTGGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTA 300
Db     241 AACCGGTATCTCTTCAACTGTGGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTA 300

Qy     301 AAGGTTGCTCGCCTGGACAACATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGC 360
Db     301 AAGGTTGTTCGCCTGGACAACATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGC 360

Qy     361 TTAAGTGGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 420
Db     361 TTAAGTGGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 420

Qy     421 CCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGA 480
Db     421 CCTCCACAGCTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGA 480

Qy     481 ATAGAACTGGCTGTGCGGCCCCACTCTGCCCCAGAATACGAGGATGAAACCATGACAGTT 540
Db     481 ATAGAACTGGCTGTGCGGCCCCACTCTGCCCCAGAATACGAGGATGAAACCATGACAGTT 540

Qy     541 TACCAGATCCCCATACACAGTGAACAGAGGAGGGGAAAGCACCAACCATGGCAGAGTCCA 600
Db     541 TACCAGATCCCCATACACAGTGAACAGAGGAGGGGAAGGCACCAACCATGGCAGAGTCCA 600

Qy     601 GAAAGGCCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCGAATGAAAAT 660
Db     601 GAAAGGCCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCGAATGAAAAT 660

Qy     661 GAGCCACACCTTCCACATGGTGTTAGCCAGAGAAGAGGGGTCAGGGACTCTTCCCTGGTC 720
Db     661 GAGCCACACCTTCCACATGGTGTTAGCCAGAGAAGAGGGGTCAGGGACTCTTCCCTGGTC 720

Qy     721 GTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGGAAACTTCTTGGTGCTCAAAGCAAAG 780
Db     721 GTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGGAAACTTCTTGGTGCTCAAAGCAAAG 780

Qy     781 GAGATGGGCCTCCCAGTTGGGACAGCTGCCATCGCTCCCATCATTGCTGCTGTCAAGGAC 840
Db     781 GAGATGGGCCTCCCAGTTGGGACAGCTGCCATCGCTCCCATCATTGCTGCTGTCAAGGAC 840

Qy     841 GGGAAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCCTCCA 900
Db     841 GGGAAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCCTCCA 900

Qy     901 GATCCTGGTGCTGCTTTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCATTCAACCCATC 960
Db     901 GATCCTGGTGCTGCTTTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCATTCAACCCATC 960

Qy     961 TGTGAGAATGCCACCTTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCGTGGCCTTGGTG 1020
Db     961 TGTGAGAATGCCACCTTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCGTGGCCTTGGTG 1020

Qy    1021 GTTCACATGGCCCCAGCATCTGTGCTTGTGGACAGCAGGTACCAGCAGTGGATGGAGAGG 1080
Db    1021 GTTCACATGGCCCCAGAATCTGTGCTTGTGGACAGCAGGTACCAGCAGTGGATGGAGAGG 1080

Qy    1081 TTTGGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTT 1140
Db    1081 TTTGGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTT 1140

Qy    1141 CGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCTC 1200
Db    1141 CGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCTC 1200

Qy    1201 ACCAGTTTCCGCTGTAAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGGTTCAGGGTGAA 1260
Db    1201 ACCAGTTTCCCCTGTAAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGGTTCAGGGTGAA 1260

Qy    1261 TGCCTCCTCAAGTACCAGCTCCGTCCCAGGAGGGAGTGGCAGAGGGATGCCATTATTACT 1320
Db    1261 TGCCTCCTCAAGTACCAGCTCCGTCCCAGGAGGGAATGGCAGAGGGATGCCATTATCACT 1320

Qy    1321 TGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGCGTG 1380
Db    1321 TGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGTGTG 1380

Qy    1381 CAGGAGTACAGGAGGAGTGCGCAGGACGGCCCAGCCCCAGCAGAGAAAAGAAGTCAGTAC 1440
Db    1381 CAGGAGTACAGGAGGAGTGTGCAGGACGTCCCAGCCCCAGCAGAGAAAAGAAGTCAGTAC 1440

Qy    1441 CCAGAAATCATCTTCCTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGT 1500
Db    1441 CCAGAAATCATCTTCCTTGGAACAGGGTCTGCCATCCCCATGAAGATTCGAAATGTCAGT 1500

Qy    1501 GCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACA 1560
Db    1501 GCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACG 1560

Qy    1561 TTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCT 1620
Db    1561 TTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCT 1620

Qy    1621 GCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGGGCTTGCCAAGTATCTTGCTG 1680
Db    1621 GCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGGGCTTGCTAAATATCTTGCTG 1680

Qy    1681 CAGAGAGAACGCGCCTTGGCATCTTTGGGAAAGCCGCTTCACCCTTTGCTGGTGGTTGCC 1740
Db    1681 CAGAGAGAACAAGCCTTGGCATCTTTGGGAAAGCCCCTTCACCCTTTGCTGGTGGTTGCC 1740

Qy    1741 CCCAACCAGCTCAAAGCCTGGCTCCAGCAGTACCACAACCAGTGCCAGGAGGTCCTGCAC 1800
Db    1741 CCCAGCCAGCTCAAAGCCTGGCTCCAGCAGTACCACAACCAGTGCCAGGAGGTCCTGCAC 1800

Qy    1801 CACATCAGTATGATTCCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCTCCAGTCCTGCA 1860
Db    1801 CACATCAGTATGATTCCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCTCCAGTCCTGCA 1860
```

```
Qy    1861 GTGGAAAGATTGATCAGTTCGCTGTTGCGAACATGTGATTTGGAAGAGTTTCAGACCTGT 1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1861 GTGGAAAGATTGATCAGTTCGCTGTTGCGAACATGTGATTTGGAAGAGTTTCAGACCTGT 1920

Qy    1921 CTGGTGCGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAA 1980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1921 CTGGTGCGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAA 1980

Qy    1981 GTGGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCTGGTCCGGATGGGGAAAGATGCC 2040
           |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    1981 GTGGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCTGGTCCGCATGGGGAAAGATGCC 2040

Qy    2041 ACCCTCCTGATACATGAAGCCACCCTGGAAGATGGTTTGGAAGAGGAAGCAGTGGAAAAG 2100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2041 ACCCTCCTGATACATGAAGCCACCCTGGAAGATGGTTTGGAAGAGGAAGCAGTGGAAAAG 2100

Qy    2101 ACACACAGCACAACGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTCATT 2160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2101 ACACACAGCACAACGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTCATT 2160

Qy    2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCCCTCTTCAGCCCCAACTTCAGC 2220
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCCCTCTTCAGCCCCAACTTCAAC 2220

Qy    2221 GAGAAAGTGGGAGTTGCCTTTGACCACATGAAGGTCTGCTTTGGAGACTTTCCAACAATG 2280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2221 GAGAAAGTGGGAGTTGCCTTTGACCACATGAAGGTCTGCTTTGGAGACTTTCCAACAATG 2280

Qy    2281 CCCAAGCTGATTCCCCCACTGAAAGCCCTGTTTGCTGGCGACATCGAGGAGATGGAGGAG 2340
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    2281 CCCAAGCTGATTCCCCCACTGAAAGCCCTGTTTGCCGGCGACATCGAGGAGATGGAGGAG 2340

Qy    2341 CGCAGGGAGAAGCGGGAGCTGCGGCAGGTGCGGGCGGCCCTCCTGTCCAGGGAGCTGGCA 2400
           |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    2341 CGCAGGGAGAAGCGGGAGCTGCGGCAGGTGCGGGCGGCCCTCCTGTCCGGGGAGCTGGCA 2400

Qy    2401 GGCGGCCTGGAGGATGGGGAGCCTCAGCAGAAGCGGGCCCACACAGAGGAGCCACAGGCC 2460
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    2401 GGCGGCCTGGAGGATGGGGAGCCTCAGCAGAAACGGGCCCACACAGAGGAGCCACAGGCC 2460

Qy    2461 AAGAAGGTCAGAGCCCAGTGA 2481
           ||||||||||||||||||||
Db    2461 AAGAAGGTCAGAGCCCAGTGA 2481
```

RESULT 9
AAC76445
ID AAC76445 standard; cDNA; 2546 BP.
XX
AC AAC76445;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF2000 polynucleotide sequence SEQ ID NO:3999.
XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.
XX
OS Homo sapiens.
XX
PN WO200058473-A2.
XX
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-US08621.
XX
PR 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2000-602362/57.
DR P-PSDB; AAB42236.
XX
PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
PS Claim 5; Page 3179-3180; 5507pp; English.
XX
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianaemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ Sequence 2546 BP; 652 A; 643 C; 686 G; 564 T; 1 other;

Query Match 71.8%; Score 1782; DB 21; Length 2546;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1813; Conservative 0; Mismatches 15; Indels 2; Gaps 2;

```
Qy     654 TGAAAATGAGCCACACCTTCCACATGGTGTTAGCCAGAGAAGAGGGGTCAGGGACTCTTC 713
           ||| ||| |         || |    ||||||||||||||||||||||||||||||||||||
Db     251 TGATAATCAAGTGTGTTTTTCCTCAGGTGTTAGCCAGAGAAGAGGGGTCAGGGACTCTTC 310

Qy     714 CCTGGTCGTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGGAAACTT-CTTGGTGCTCA 772
           |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db     311 CCTGGTCGTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGGAAACTTACTTGGTGCTCA 370

Qy     773 AAGCAAAGGAGATGGGCCTCCCAGTTGGGACAGCTGCCATCGCTCCCATCATTGCTGCTG 832
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     371 AAGCAAAGGAGATGGGCCTCCCAGTTGGGACAGCTGCCATCGCTCCCATCATTGCTGCTG 430

Qy     833 TCAAGGACGGGAAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTA 892
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     431 TCAAGGACGGGAAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTA 490

Qy     893 CTCCTCCAGATCCTGGTGCTGCTTTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCATTC 952
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     491 CTCCTCCAGATCCTGGTGCTGCTTTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCATTC 550

Qy     953 AACCCATCTGTGAGAATGCCACCTTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCGTGG 1012
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    551 AACCCATCTGTGAGAATGCCACCTTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCGTGG 610

Qy   1013 CCTTGGTGGTTCACATGGCCCCAGCATCTGTGCTTGTGGACAGCAGGTACCAGCAGTGGA 1072
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    611 CCTTGGTGGTTCACATGGCCCCAGCATCTGTGCTTGTGGACAGCAGGTACCAGCAGTGGA 670

Qy   1073 TGGAGAGGTTTGGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTC 1132
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    671 TGGAGAGGTTTGGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTC 730

Qy   1133 ACAACCTTCGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCC 1192
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    731 ACAACCTTCGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCC 790

Qy   1193 CCCTGCTCACCAGTTTCCGCTGTAAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGGTTC 1252
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    791 CCCTGCTCACCAGTTTCCGCTGTAAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGGTTC 850

Qy   1253 AGGGTGAATGCCTCCTCAAGTACCAGCTCCGTCCCAGGAGGGAGTGGCAGAGGGATGCCA 1312
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    851 AGGGTGAATGCCTCCTCAAGTACCAGCTCCGTCCCAGGAGGGAGTGGCAGAGGGATGCCA 910

Qy   1313 TTATTACTTGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGC 1372
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    911 TTATTACTTGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGC 970

Qy   1373 AGAGCGTGCAGGAGTACAGGAGGAGTGCGCAGGACGGCCCAGCCCCAGCAGAGAAAAGAA 1432
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    971 AGAGCGTGCAGGAGTACAGGAGGAGTGCGCAGGACGGCCCAGCCCCAGCAGAGAAAAGAA 1030

Qy   1433 GTCAGTACCCAGAAATCATCTTCCTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAA 1492
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1031 GTCAGTACCCAGAAATCATCTTCCTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAA 1090

Qy   1493 ATGTCAGTGCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGACTGTGGTG 1552
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1091 ATGTCAGTGCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGACTGTGGTG 1150

Qy   1553 AGGGCACATTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCA 1612
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1151 AGGGCACATTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCA 1210

Qy   1613 CCCTGGCTGCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGGGCTTGCCAAGTA 1672
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1211 CCCTGGCTGCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGGGCTTGCCAAGTA 1270

Qy   1673 TCTTGCTGCAGAGAGAACGCGCCTTGGCATCTTTGGGAAAGCCGCTTCACCCTTTGCTGG 1732
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1271 TCTTGCTGCAGAGAGAACGCGCCTTGGCATCTTTGGGAAAGCCGCTTCACCCTTTGCTGG 1330

Qy   1733 TGGTTGCCCCCAACCAGCTCAAAGCCTGGCTCCAGCAGTACCACAACCAGTGCCAGGAGG 1792
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1331 TGGTTGCCCCCAACCAGCTCAAAGCCTGGCTCCAGCAGTACCACAACCAGTGCCAGGAGG 1390

Qy   1793 TCCTGCACCACATCAGTATGATTCCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCTCCA 1852
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1391 TCCTGCACCACATCAGTATGATTCCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCTCCA 1450

Qy   1853 GTCCTGCAGTGGAAAGATTGATCAGTTCGCTGTTGCGAACATGTGATTTGGAAGAGTTTC 1912
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1451 GTCCTGCAGTGGAAAGATTGATCAGTTCGCTGTTGCGAACATGTGATTTGGAAGAGTTTC 1510

Qy   1913 AGACCTGTCTGGTGCGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTG 1972
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1511 AGACCTGTCTGGTGCGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTG 1570

Qy   1973 GCTGGAAAGTGGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCTGGTCCGGATGGGGA 2032
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1571 GCTGGAAAGTGGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCTGGTCCGGATGGGGA 1630

Qy   2033 AAGATGCCACCCTCCTGATACATGAAGCCACCCTGGAAGATGGTTTGGAAGAGGAAGCAG 2092
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1631 AAGATGCCACCCTCCTGATACATGAAGCCACCCTGGAAGATGGTTTGGAAGAGGAAGCAG 1690

Qy   2093 TGGAAAAGACACACAGCACAACGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGG 2152
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1691 TGGAAAAGACACACAGCACAACGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGG 1750

Qy   2153 AGTTCATTATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCCCTCTTCAGCCCCA 2212
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1751 AGTTCATTATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCCCTCTTCAGCCCCA 1810

Qy   2213 ACTTCAGCGAGAAAGTGGGAGTTGCCTTTGACCACATGAAGGTCTGCTTTGGAGACTTTC 2272
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1811 ACTTCAGCGAGAAAGTGGGAGTTGCCTTTGACCACATGAAGGTCTGCTTTGGAGACTTTC 1870

Qy   2273 CAACAATGCCCAAGCTGATT-CCCCCACTGAAAGCCCTGTTTGCTGGCGACATCGAGGAG 2331
          |||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db   1871 CAACAATGCCCAAGCTGATTCCCCCCACTGAAAGCCCTGTTTGCTGGCGACATCGAGGAG 1930

Qy   2332 ATGGAGGAGCGCAGGGAGAAGCGGGAGCTGCGGCAGGTGCGGGCGGCCCTCCTGTCCAGG 2391
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1931 ATGGAGGAGCGCAGGGAGAAGCGGGAGCTGCGGCAGGTGCGGGCGGCCCTCCTGTCCAGG 1990

Qy   2392 GAGCTGGCAGGCGGCCTGGAGGATGGGGAGCCTCAGCAGAAGCGGGCCCACACAGAGGAG 2451
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1991 GAGCTGGCAGGCGGCCTGGAGGATGGGGAGCCTCAGCAGAAGCGGGCCCACACAGAGGAG 2050

Qy   2452 CCACAGGCCAAGAAGGTCAGAGCCCAGTGA 2481
          |||||||||||||||||||||||||||||
Db   2051 CCACAGGCCAAGAAGGTCAGAGCCCAGTGA 2080
```

RESULT 10
AAS99131
ID AAS99131 standard; cDNA; 2470 BP.
XX
AC AAS99131;
XX
DT 12-MAR-2002 (first entry)
XX
DE Mouse ELAC2 cDNA.
XX
KW Human; mouse; HPC2; prostate cancer; neoplastic growth; cytostatic; ss;
KW gene therapy; prostate cancer predisposing gene; chimpanzee; gorilla;
KW sequencing primer; PCR primer.
XX
OS Mus musculus.
XX
PN WO200185911-A2.
XX
PD 15-NOV-2001.
XX
PF 07-MAY-2001; 2001WO-US14602.
XX
PR 05-MAY-2000; 2000US-0564805.
XX
PA (MYRI-) MYRIAD GENETICS INC.
PA (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI Tavtigian SV, Teng DHF, Simard J, Rommens JM;
XX
DR WPI; 2002-066599/09.
DR P-PSDB; AAU73591.
XX
PT Novel nucleic acid sequence encoding HPC2 polypeptide, which is marker
PT for prostate cancer, is useful in gene therapy techniques to restore
PT HPC2 normal levels by which neoplastic growth is suppressed in
PT recipient cell -
XX
PS Claim 82; Page 192-195; 239pp; English.
XX
CC The invention relates to a human prostate cancer predisposing gene coding
CC for an HPC2 polypeptide. The DNA and protein sequences are useful as
CC diagnostic reagents for identifying a mutant HPC2 nucleotide sequence in
CC a suspected mutant HPC2 allele by comparing the sequence of the suspected

CC mutant HPC2 allele with a wild-type HPC2 sequence. The sequences are also
CC useful for detecting an alteration in HPC2, where the alteration is
CC associated with cancer in a human. The method involves analysing an HPC2
CC gene or an HPC2 gene expression product from a tissue of the human. The
CC HPC2 gene is useful as a marker for prostate cancer and can be used in
CC gene therapy techniques to suppress neoplastic growth of recipient cells
CC which carry the mutant HPC2 allele. The sequences represent primers used
CC in the methods of the invention, cDNA encoding human and mouse HPC2 and
CC cDNA encoding HPC2 paralogues and orthologues.
XX
SQ Sequence 2470 BP; 614 A; 664 C; 679 G; 510 T; 3 other;

Query Match 66.3%; Score 1645.6; DB 24; Length 2470;
Best Local Similarity 81.6%; Pred. No. 0;
Matches 1958; Conservative 0; Mismatches 417; Indels 24; Gaps 4;

```
Qy      58 CGCACCATATCGCAGGCACCCGCCCGCCGCGAGCGGCCGCGCAAGGACCCGCTGCGGCAC 117
           ||||||||| |||||||   | || || ||   |||||| | ||| ||||| ||||| |||
Db      40 CGCACCATGTCGCAGGGTTCGGCTCGTCGGCCGCGGCCACCCAAAGACCCACTGCGACAC 99

Qy     118 CTGCGCACGCGAGAGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAAACACCGTGTAC 177
           ||||| ||||| |||||||||||| ||      | |  |||| ||||| |||||||||||
Db     100 CTGCGTACGCGGGAGAAGCGCGGCCC------GGGTCCCGGGGGCCCGAACACCGTGTAC 153

Qy     178 CTGCAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAG 237
           |||||||||||||| |||||  |||||||| |||| || || ||||| |||||||| ||
Db     154 CTGCAGGTGGTGGCGGCGGGCGGCCGGGACGCGGGGGCTGCTCTCTATGTCTTCTCGGAA 213

Qy     238 TTCAACCGGTATCTCTTCAACTGTGGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAG 297
           | |||| |||| || || ||||| |||||||||| ||  |||| |||||||| ||||||
Db     214 TACAACAGGTACCTTTTTAACTGCGGAGAAGGCGTCCAACGACTTATGCAGGAACACAAG 273

Qy     298 TTAAAGGTTGCTCGCCTGGACAACATATTCCTGACACGAATGCACTGGTCTAATGTTGGG 357
             |   |       || |||||||| || ||||| || ||||| ||||| |||||||||
Db     274 ACTGAAAGTCGCTCGCTTGACAACATCTTTCTGACTCGGATGCATTGGTCAAATGTTGGG 333

Qy     358 GGCTTAAGTGGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCT 417
           || ||  |||||||||| | ||||||||||||||||||||||| || ||||| || |||
Db     334 GGGTTGTGTGGAATGATTTTAACTTTAAAGGAAACCGGGCTTCCCAAATGTGTTCTGTCT 393

Qy     418 GGACCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAA 477
           ||||| ||||| ||||| ||||| || ||||||||||||||||||||||||||||||||||
Db     394 GGACCACCACAGCTGGAGAAATATCTAGAAGCAATCAAAATATTTTCTGGTCCATTGAAA 453

Qy     478 GGAATAGAACTGGCTGTGCGGCCCCACTCTGCCCCAGAATACGAGGATGAAACCATGACA 537
           |||||||||||||| |||||||| |||||||| |||||||| ||||||| ||||||||
Db     454 GGAATAGAACTGGCCGTGCGGCCTCACTCTGCACCAGAATACAAGGATGAGACCATGACT 513

Qy     538 GTTTACCAGATCCCCATACACAGTGAACAGAGGAGGGGAAAGCACCAACCATGGCAGAGT 597
           |||||||| |||| || |||||||||| |||| | |||||||| || ||||  |||||
Db     514 GTTTACCAGGTCCCTATCCACAGTGAACGGAGGTGTGGAAAGCAACAGCCATCCCAGAGC 573

Qy     598 CCAGAAAGGCCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCGAATGAA 657
           ||   ||   ||| || ||||||||||||  | |  || || ||||| |  ||   ||||
Db     574 CCCAGAACATCTCCCAACAGGCTCAGTCCCAAACAGTCATCGGACTCTGGATCAGCTGAA 633

Qy     658 AATGAGCCACACCTTCCACATGGTGTTAGCCAGAGAAGAGGGGTCAGGGACTCTTCCCTG 717
           |||| ||               ||| | |  || |  | ||| |||||||  | ||| |
Db     634 AATGGGC-------------AGTGCCAACAGGAAAGCATGGGGCAGGGAC-CCTCCTTA 678

Qy     718 GTCGTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGGAAACTTCTTGGTGCTCAAAGCA 777
           || ||||||||  |||| |||||||||| | || ||||||||||||||||| ||||||
Db     679 GTGGTAGCTTTTGTCTGCAAGCTTCACTTGAGGAAAGGAAACTTCTTGGTGCTTAAAGCA 738

Qy     778 AAGGAGATGGGCCTCCCAGTTGGGACAGCTGCCATCGCTCCCATCATTGCTGCTGTCAAG 837
           |||||| ||||||| || |||||||| || ||||| || ||||||||||||||||||||
Db     739 AAGGAGCTGGGCCTTCCTGTTGGGACGGCCGCCATTGCACCCATCATTGCTGCTGTCAAG 798

Qy     838 GACGGGAAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCCT 897
           |||||||| || |||||| | ||||||||||||||||   |||||||||| |||| ||
Db     799 GACGGGAAGAGTATCACTTACGAAGGAAGAGAGATTGCTGCTGAAGAGCTTTGTACACCC 858

Qy     898 CCAGATCCTGGTGCTGCTTTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCATTCAACCC 957
           |||||||||||  ||  ||  | |||||||| ||||| |||||| | ||||| |  |||
Db     859 CCAGATCCTGGTCTTGTATTCATCGTGGTAGAGTGTCCTGATGAAGGATTCATCCTGCCC 918

Qy     958 ATCTGTGAGAATGCCACCTTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCGTGGCCTTG 1017
           |||||||||| | ||||||| | |||||||| | | |||| ||||| || |||||  ||
Db     919 ATCTGTGAGAACGACACCTTTAAAAGGTACCAGGCAGAGGCTGATGCACCTGTGGCGCTG 978

Qy    1018 GTGGTTCACATGGCCCCAGCATCTGTGCTTGTGGACAGCAGGTACCAGCAGTGGATGGAG 1077
           ||||| ||||| ||||||| |||||| ||  | |||||||| ||||||||||||||||||
Db     979 GTGGTCCACATAGCCCCAGAATCTGTACTCATCGACAGCAGATACCAGCAGTGGATGGAG 1038

Qy    1078 AGGTTTGGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTCACAAC 1137
           ||||| |||||||||||| |||||| || | |||||||||| ||  |||| || ||||||
Db     1039 AGGTTCGGGCCTGACACACAGCACCTGATTCTGAATGAGAATTGCCCCTCGGTCCACAAC 1098

Qy    1138 CTTCGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTG 1197
           || |||||||||||||||| |||||||||| |||||||||||| ||||||||||||| |
Db     1099 CTGCGCAGCCACAAGATTCAGACCCAGCTCAGCCTCATCCACCCTGACATCTTCCCCCAG 1158

Qy    1198 CTCACCAGTTTCCGCTGTAAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGGTTCAGGGT 1257
           || ||||| |||    ||||| |||| ||  |||||||||| ||||| |  |||| ||||
Db     1159 CTTACCAGCTTCTATAGTAAGGAGGAAGGGTCCACCCTCAGCGTGCCAACAGTTCGGGGT 1218

Qy    1258 GAATGCCTCCTCAAGTACCAGCTCCGTCCCAGGAGGGAGTGGCAGAGGGATGCCATTATT 1317
           ||||||||||||||||     |||| |||| ||| |||||||||||||| |||   |
Db     1219 GAATGCCTCCTCAAGTATTCAGTCCGCCCCAAGAGAGAGTGGCAGAGGGATACCACACTC 1278

Qy    1318 ACTTGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGC 1377
             |||||| |||| |||||||||| ||||||  || |||| ||||  |||||| ||||
Db     1279 GACTGCAATACTGATGAATTCATAGCTGAGGCCTTGGAGCTCCCCAGTTTCCAGGAGAGT 1338

Qy    1378 GTGCAGGAGTACAGGAGGAGTGCGCAGGACGGCCCAGCCCCAGCAGAGAAAAGAAGTCAG 1437
           ||| |||||||  ||| ||  | ||||||   ||||||||||||||||||||||| |||
Db     1339 GTGGAGGAGTATCGGAAGAACGTGCAGGAAAACCCAGCCCCAGCAGAGAAAAGAAGCCAG 1398

Qy    1438 TACCCAGAAATCATCTTCCTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTC 1497
           || || |||||  ||||||| || || |||||||||||||| ||| |||| ||||||||||
Db     1399 TATCCTGAAATTGTCTTCCTGGGTACGGGGTCTGCCATCCCAATGGAGATCCGAAATGTC 1458

Qy    1498 AGTGCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGC 1557
           ||| ||||||| |||||| ||||||| |||| ||  |||| |||| |||| || |||
Db     1459 AGTTCCACACTCGTCAACCTAAGCCCTGACAAGTCAGTGCTCCTGGATTGTGGAGAAGGC 1518

Qy    1558 ACATTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTG 1617
           || |||||||| |||||||||||||||| | ||  | ||| | ||| |  ||| |||
Db     1519 ACTTTTGGGCAGTTGTGCCGTCATTACGGACAGCAAATAGACCGAGTCTTATGCAGCCTC 1578

Qy    1618 GCTGCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGGGCTTGCCAAGTATCTTG 1677
           | |||||||||||||||||||||||||||||| ||||||||||||||  | |||||||
Db     1579 ACGGCTGTGTTTGTGTCCCACCTGCACGCCGACCACCACACGGGCTTGCTGAATATCTTG 1638

Qy    1678 CTGCAGAGAGAACGCGCCTTGGCATCTTTGGGAAAGCCGCTTCACCCTTTGCTGGTGGTT 1737
           |||||||||| |  || |||||||| |||| || ||  | || || |||||  ||||
Db     1639 CTGCAGAGAGAGCATGCGTTGGCATCTCTGGGGAAACCCTTCCAGCCCTTGCTTGTGGTG 1698

Qy    1738 GCCCCCAACCAGCTCAAAGCCTGGCTCCAGCAGTACCACAACCAGTGCCAGGAGGTCCTG 1797
           || || | ||||||||  |||||||| |||||||| |||||||| ||||||||  | |||
Db     1699 GCTCCTACCCAGCTCAGGGCCTGGCTGCAGCAGTATCACAACCACTGCCAGGAGATTCTG 1758

Qy    1798 CACCACATCAGTATGATTCCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCTCCAGTCCT 1857
           |||||| |||||||||||||||||||||||||||||||||||||| |||||| ||| |||||| | ||
Db     1759 CACCACGTCAGTATGATTCCTGCCAAATGCCTTCAGAAAGGGGCAGAGGTCTCCAATACT 1818

Qy    1858 GCAGTGGAAAGATTGATCAGTTCGCTGTTGCGAACATGTGATTTGGAAGAGTTTCAGACC 1917
             || |||||||  |||| || | |||||||  |||||||| || ||||| |||||||||
Db     1819 ACATTGGAAAGGCTGATAAGCTTGCTGTTGGAAACATGTGACTTAGAAGAATTTCAGACC 1878

Qy    1918 TGTCTGGTGCGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGG 1977
           || ||||| |||||||||||||||| |||||||||| ||||| ||  | |||||||||
Db     1879 TGCCTGGTACGGCACTGCAAGCATGCTTTTGGCTGTGCACTGGTACATTCATCTGGCTGG 1938

Qy    1978 AAAGTGGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCTGGTCCGGATGGGGAAAGAT 2037
           ||||| ||||| || ||||| |||||||||| ||||||||||| |||||||||||||
```

```
Db   1939 AAAGTCGTCTACTCGGGGGATACCATGCCCTGTGAGGCTCTGGTCCAGATGGGGAAAGAT 1998

Qy   2038 GCCACCCTCCTGATACATGAAGCCACCCTGGAAGATGGTTTGGAAGAGGAAGCAGTGGAA 2097
          |||||||||||||||||||||||||||| ||||| |||   ||||||||||||||||| ||
Db   1999 GCCACCCTCCTGATACATGAAGCCACTCTGGAGGATCNCTTGGAAGAGGAAGCAGTAGAG 2058

Qy   2098 AAGACACACAGCACAACGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTC 2157
          | |||||||||||| || ||||| || || |  |||||||||||||||| |||||||||
Db   2059 AGGACACACAGCACCACCTCCCAGGCTATTAATGTGGGGATGCGGATGAATGCGGAGTTC 2118

Qy   2158 ATTATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCCCTCTTCAGCCCCAACTTC 2217
          || |||||||||||||||| ||||| || || ||| ||||||| |||||||| |||||
Db   2119 ATCATGCTGAACCACTTCAGTCAGCGGTACGCNAAGATCCCCCTTTTCAGCCCTGACTTC 2178

Qy   2218 AGCGAGAAAGTGGGAGTTGCCTTTGACCACATGAAGGTCTGCTTTGGAGACTTTCCAACA 2277
          | |||||||| ||  | ||||||||||||||||||||| |||||||||| || |||
Db   2179 AACGAGAAAGTTGGCATCGCCTTTGACCACATGAAGGTCTGNTTTGGAGACTTCCCGACA 2238

Qy   2278 ATGCCCAAGCTGATTCCCCCACTGAAAGCCCTGTTTGCTGGCGACATCGAGGAGATGGAG 2337
          ||||||||||||||||||||||||| |||||||||| || ||||| || ||||||| |
Db   2239 GTGCCCAAGCTGATTCCCCCACTGAAGGCCCTGTTTGCAGGTGACATTGAAGAGATGGTG 2298

Qy   2338 GAGCGCAGGGAGAAGCGGGAGCTGCGGCAGGTGCGGGCGGCCCTCCTGTCCAGGGAGCTG 2397
          || |||||||||||| ||||||| |||| |||||| || ||||||||| ||    ||| |
Db   2299 GAACGCAGGGAGAAGAGGGAGCTACGGCTGGTGCGAGCAGCCCTCCTGACC---CAGCAG 2355

Qy   2398 GCAGGCGGCCTGGAGGATGGGGAGCCTCAGCAGAAGCGGGCCCACACAGAGGAGCCACA 2456
          |||| | |||  |||||  | || || || ||||||||||||||||| || |||||
Db   2356 GCAGACAGCCCAGAGGACAGAGAACCCCAACAGAAGCGGGCCCACACAGATGAACCACA 2414
```

RESULT 11
AAS72207
ID AAS72207 standard; cDNA; 1402 BP.
XX
AC AAS72207;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #8011.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG08020.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 8011; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1402 BP; 338 A; 371 C; 377 G; 316 T; 0 other;

Query Match 43.6%; Score 1080.8; DB 23; Length 1402;
Best Local Similarity 98.7%; Pred. No. 1.2e-265;
Matches 1173; Conservative 0; Mismatches 7; Indels 8; Gaps 8;

```
Qy    865 AGAGAGATTTTGGCTGAAGAGCTGTGTACTCCTCCAGATCCTGGTGCTGCTTTTGTGGTG 924
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     17 AGGGAGATTTTGGCTGAAGAGCTGTGTACTCCTCCAGATCCTGGTGCTGCTTTTGTGGTG 76

Qy    925 GTAGAATGTCCAGATGAAAGCTTCATTCAACCCATCTGTGAGAATGCCACCTTTCAGAGG 984
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     77 GTAGAATGTCCAGATGAAAGCTTCATTCAACCCATCTGTGAGAATGCCACCTTTCAGAGG 136

Qy    985 TACCAAGGAAAGGCAGATGCCCCCGTGGCCTTGGTGGTTCACATGGCCCCAGCATCTGTG 1044
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    137 TACCAAGGAAAGGCAGATGCCCCCGTGGCCTTGGTGGTTCACATGGCCCCAGCATCTGTG 196

Qy   1045 CTTGTGGACAGCAGGTACCAGCAGTGGATGGAGAGGTTTGGGCCTGACACCCAGCACTTG 1104
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    197 CTTGTGGACAGCAGGTACCAGCAGTGGATGGAGAGGTTTGGGCCTGACACCCAGCACTTG 256

Qy   1105 GTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTTCGCAGCCACAAGATTCAAACCCAG 1164
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    257 GTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTTCGCAGCCACAAGATTCAAACCCAG 316

Qy   1165 CTCAACCTCATCCACCCGGACATCTTCCCCCTGCTCACCAGTTTCCGCTGTAAGAAGGAG 1224
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    317 CTCAACCTCATCCACCCGGACATCTTCCCCCTGCTCACCAGTTTCCGCTGTAAGAAGGAG 376

Qy   1225 GGCCCCACCCTCAGTGTGCCCATGGTTCAGGGTGAATGCCTCCTCAAGTACCAGCTCCGT 1284
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    377 GGCCCCACCCTCAGTGTGCCCATGGTTCAGGGTGAATGCCTCCTCAAGTACCAGCTCCGT 436

Qy   1285 CCCAGGAGGGAGTGGCAGAGGGATGCCATTATTACTTGCAATCCTGAGGAATTCATAGTT 1344
           |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    437 TCCAGGAGGGAGTGCCAGAGGGATGCCATTATTACTTGCAATCCTGAGGAATTCATAGTT 496

Qy   1345 GAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGCG-TGCAGGAGTACAGGAGGAGTGCGCA 1403
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    497 GAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGCGTTGCAGGAGTACAGGAGGAGTGCGCA 556

Qy   1404 GGACGGCCCAGCCCCAGCAGAGAAAAGAAGTCAGTACCCAGAAATCATCTTCCTTGGAAC 1463
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    557 GGACGGCCCAGCCCCAGCAGAGAAAAGAAGTCAGTACCCAGAAATCATCTTCCTTGGAAC 616

Qy   1464 AGGGTCTGCCATCCCGATGAAGATTCG-AAATGTCAGTGCCACACTTGTC-AACATAAGC 1521
          |||||||||||||||||||||||||| |||||||||||||||||||| |||||||||
Db    617 AGGGTCTGCCATCCCGATGAAGATTCGAAAATGTCAGTGCCACACTTGTCAAACATAAGC 676

Qy   1522 CCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACATTTGGGCAGCTGTGCCGTCAT 1581
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    677 CCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACATTTGGGCAGCTGTGCCGTCAT 736

Qy   1582 TACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCTGCTGTGTTTGTGTCCCACCTG 1641
```

```
Db    737 TACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCTGCTGTGTTTGTGTCCCACCTG 796

Qy   1642 CACGCAGATCACCACACGGGCTTGCCAAGTATCTTGCTGCAGAGAGAACGCGCCTTGGCA 1701
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    797 CACGCAGATCACCACACGGGCTTGCCAAGTATCTTGCTGCAGAGAGAACGCGCCTTGGCA 856

Qy   1702 TCTTTGGGAAAGCCGCTTCACCCTTTGCTGGTGGTTGCCCCCAACCAGCTCAAAGCCTGG 1761
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    857 TCTTTGGGAAAGCCGCTTCACCCTTTGCTGGTGGTTGCCCCCAACCAGCTCAAAGCCTGG 916

Qy   1762 CTCCAGCAGTACCACAACCAGTGCCAGGAGGTCCTGCACCACATCAGTATGATTCCTGCC 1821
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    917 CTCCAGCAGTACCACAACCAGTGCCAGGAGGTCCTGCACCACATCAGTATGATTCCTGCC 976

Qy   1822 AAA-TGCCTTCAGGAAGG-GGCTGAGATCTCCAGTCCTGCAGTGGAAAGATTGA-TCAGT 1878
          ||| |||||||||||||| |||||||||||||||||||||||||||||||||| |||||
Db    977 AAATTGCCTTCAGGAAGGTGGCTGAGATCTCCAGTCCTGCAGTGGAAAGATTGATTCAGT 1036

Qy   1879 TCGCTGTTGCGAACATGTGATTT-GGAAGAGTTTCAGACCTGTCTGGTGCGGCACTGCAA 1937
          |||||||||||||||||||||| |||||||||||| |||||| |||||||||| |||||
Db   1037 TCGCTGTTGCGAACATGTGATTTGGGAAGAGTTTCAAACCTGTTTGGTGCGGCATTGCAA 1096

Qy   1938 GCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTGGC-TGGAAAGTGGTCTATTCCGGGG 1996
          |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db   1097 GCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTTGGAAAGTGGTCTATTCCGGGG 1156

Qy   1997 ACACCATGCCCTGCGAGGCTCTGGTCCGGATGGGGAAAGATGCCACCC 2044
          |||||||||||||||||||||||||||||||||||||||||||||| |
Db   1157 ACACCATGCCCTGCGAGGCTCTGGTCCGGATGGGGAAAGATGCCACAC 1204
```

RESULT 12
AAS72208
ID AAS72208 standard; cDNA; 1450 BP.
XX
AC AAS72208;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #8012.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG08021.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 8012; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1450 BP; 355 A; 382 C; 418 G; 294 T; 1 other;

Query Match 20.9%; Score 518.2; DB 23; Length 1450;
Best Local Similarity 90.3%; Pred. No. 5.4e-122;
Matches 670; Conservative 0; Mismatches 49; Indels 23; Gaps 10;

```
Qy   1763 TCCAGCAGTACCACAACCAGTGCCAGGAGGTCCTGCACCACATCAGTATGATTCCTGCCA 1822
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    141 TCCAGCAGTACCACAACCAGTGCCAGGAGGTCCTGCACCACATCAGTATGATTCCTGCCA 200

Qy   1823 AATGCCTTCAGGAAGGGGCTGAGATCTCCAGTCCTGCAGTGGAAAGATTGATCAGTTCGC 1882
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    201 AATGCCTTCAGGAAGGGGCTGAGATCTCCAGTCCTGCAGTGGAAAGATTGATCAGTTCGC 260

Qy   1883 TGTT-GCGAACATGTGATTTGGAAGAGTTTCAGACCTGTCTGGTGCGGCACTGCAAGCAT 1941
          |||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    261 TGTTGGACAACATGTGATTTGGAAGAGTTTCAGACCTGTCTGGTGCGGCACTGCAAGCAT 320

Qy   1942 GCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAAGTGGTCTATTCCGGGGACACC 2001
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    321 GCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAAGTGGTCTATTCCGGGGACACC 380

Qy   2002 ATGCCCTGCGAGGC-TCTGGTCCGGATGGGGAAAGATGCCACCCTCCTGATACATGAAGC 2060
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    381 ATGCCCTGCGAGGCTTCTGGTCCGGATGGGGAAAGATGCCACCCTCCTGATACATGAAGC 440

Qy   2061 CACCCTGGAAGATGGTTTGGAAGAGGAAGCA-----GTGGAAAAGACACACAGCACAACG 2115
          |||||||| |    | ||    || || | |      ||||||||||||||||||||||
Db    441 CACCCTGGGAAGATGGTTTTNGGAAGAGGGAAGCAGTTGGAAAAGACACACAGCACAACG 500

Qy   2116 TCCCAAGCCATCAGCG---TGGGGATGCGGATGAACGCGGAGTT--CATTATGCTGAACC 2170
          ||||||| |    |    ||||||||||||||||||||||||    |   | ||||||
Db    501 TCCCAAGGCCATTCAGCGTGGGGGATGCGGATGAACGCGGAGTTTCATTATGGTTGAACC 560

Qy   2171 ACTTCAGCCAGCGCTA----TGCCAAGGTCCCCCTCTTCAGCCCCAACTTCAGCGAGAAA 2226
          ||||        |       ||||||||||||||||||||||||||||||||||||||||
Db    561 ACTTTCAGGCCAGGCGCTATTGCCAAGGTCCCCCTCTTCAGCCCCAACTTCAGCGAGAAA 620

Qy   2227 GTGGGAGTTGCCTTTGACCACATGAAGGTCTGCTTTGGAGACTTTCCAACAATGCCCAAG 2286
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    621 GTGGGAGTTGCCTTTGACCACATGAAGGTCTGCTTTGGAGACTTTCCAACAATGCCCAAG 680

Qy   2287 CTGATTCCCCCACTGAAAGCCCTGTTTGCTGGCGACATCGAGGAGAT-GGAGGAGCGCAG 2345
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    681 CTGATTCCCCCACTGAAAGCCCTGTTTGCTGGCGACATCGAGGAGATGGGAGGAGCGCAG 740

Qy   2346 GGAGAAGCGGGAGCTGCGG--CAGGTGCGGGCGGCCCTCCTGT--CCAGGGAGCTGGCAG 2401
          |||||||||||||||||    |    |||||||||||||||  |  ||||||||||||
Db    741 GGAGAAGCGGGAGCTGCGGGCAGGTGCGGGGCGGCCCTCCTGTTCCAGGGGAGCTGGCAG 800

Qy   2402 GCGGCC--TGGAGGATGGGGAGCCTCAGCAGAAGCGGGCCCACACAGAGGAGCCACAGGC 2459
          ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    801 GCGGCCTTTGGAGGATGGGGAGCCTCAGCAGAAGCGGGCCCACACAGAGGAGCCACAGGC 860
```

```
Qy   2460 CAAGAAGGTCAGAGCCCAGTGA 2481
          ||||||||||||||||||||||
Db    861 CAAGAAGGTCAGAGCCCAGTGA 882


RESULT 13
AAH05835
ID   AAH05835 standard; cDNA; 584 BP.
XX
AC   AAH05835;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA clone (5'-primer) SEQ ID NO:2670.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;   99JP-0248036.
PR   27-AUG-1999;   99JP-0300253.
PR   11-JAN-2000; 2000JP-0118776.
PR   02-MAY-2000; 2000JP-0183767.
PR   09-JUN-2000; 2000JP-0241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602
PT   full-length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full-length cDNAs -
XX
PS   Claim 1; SEQ ID 2670; 2537pp + CD ROM; English.
XX
CC   The present invention describes primer sets for synthesising 5602
CC   full-length cDNAs defined in the specification. Where a primer set
CC   comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC   to the complementary strand of a polynucleotide which comprises one of
CC   the 5602 nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC   AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 584 BP; 122 A; 166 C; 178 G; 115 T; 3 other;

  Query Match             19.2%;  Score 475.8;  DB 22;  Length 584;
  Best Local Similarity   98.0%;  Pred. No. 2.4e-111;
  Matches  491;  Conservative    0;  Mismatches    9;  Indels    1;  Gaps    1;

Qy      1 ATGTGGGCGCTTTGCTCGCTGCTGCGGTCCGCGGCCGGACGCACCATGTCGCAGGGACGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     75 ATGTGGGCGCTTTGCTCGCTGCTGCGGTCCGCGGCCGGACGCACCATGTCGCAGGGACGC 134

Qy     61 ACCATATCGCAGGCACCCGCCCGCCGCGAGCGGCCGCGCAAGGACCCGCTGCGGCACCTG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    135 ACCATATCGCAGGCACCCGCCCGCCGCGAGCGGCCGCGCAAGGACCCGCTGCGGCACCTG 194

Qy    121 CGCACGCGAGAGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAAACACCGTGTACCTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    195 CGCACGCGAGAGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAAACACCGTGTACCTG 254

Qy    181 CAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    255 CAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTC 314

Qy    241 AACCGGTATCTCTTCAACTGTGGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    315 AACCGGTATCTCTTCAACTGTGGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTA 374

Qy    301 AAGGTTGCTCGCCTGGACAACATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    375 AAGGTTGCTCGCCTGGACAACATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGC 434

Qy    361 TTAAGTGGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 420
          |||| |||||||||||||||||||| |||||| ||||||| |||||||| |||||||||
Db    435 TTAAATGGAATGATTCTTACTTTAAAAGAAACC-GGCTTCCCAAGTGTGTGCTTTCTGGA 493

Qy    421 CCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGA 480
          |||||||||||||||||||||||||| ||||||||||||||||||||| |||||| ||
Db    494 CCTCCACAACTGGAAAAATACCTCGAAACAATCAAAATATTTTCTGGTCCNTTGAAAAGA 553

Qy    481 ATAGAACTGGCTGTGCGGCCC 501
          ||| |||||||| ||||||||
Db    554 ATANAACTGGCTATGCGGCCC 574


RESULT 14
AAS99124
ID   AAS99124 standard; DNA; 350 BP.
XX
AC   AAS99124;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Human prostate cancer predisposing gene (HPC2) DNA full length exon #1.
XX
KW   Human; mouse; HPC2; prostate cancer; neoplastic growth; cytostatic; ds;
KW   gene therapy; prostate cancer predisposing gene.
XX
OS   Homo sapiens.
XX
PN   WO200185911-A2.
XX
PD   15-NOV-2001.
XX
PF   07-MAY-2001; 2001WO-US14602.
XX
PR   05-MAY-2000; 2000US-0564805.
XX
PA   (MYRI-) MYRIAD GENETICS INC.
PA   (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI   Tavtigian SV,  Teng DHF,  Simard J,  Rommens JM;
XX
DR   WPI; 2002-066599/09.
DR   P-PSDB; AAU73587.
XX
PT   Novel nucleic acid sequence encoding HPC2 polypeptide, which is marker
PT   for prostate cancer, is useful in gene therapy techniques to restore
PT   HPC2 normal levels by which neoplastic growth is suppressed in
PT   recipient cell  -
```

```
XX
PS   Example 5; Fig 4; 239pp; English.
XX
CC   The invention relates to a human prostate cancer predisposing gene coding
CC   for an HPC2 polypeptide. The DNA and protein sequences are useful as
CC   diagnostic reagents for identifying a mutant HPC2 nucleotide sequence in
CC   a suspected mutant HPC2 allele by comparing the sequence of the suspected
CC   mutant HPC2 allele with a wild-type HPC2 sequence. The sequences are also
CC   useful for detecting an alteration in HPC2, where the alteration is
CC   associated with cancer in a human. The method involves analysing an HPC2
CC   gene or an HPC2 gene expression product from a tissue of the human. The
CC   HPC2 gene is useful as a marker for prostate cancer and can be used in
CC   gene therapy techniques to suppress neoplastic growth of recipient cells
CC   which carry the mutant HPC2 allele. The sequences represent DNA encoding
CC   human and mouse HPC2 and fragments of HPC2.
XX
SQ   Sequence 350 BP; 47 A; 125 C; 126 G; 52 T; 0 other;

  Query Match             10.0%;  Score 247.4;  DB 24;  Length 350;
  Best Local Similarity   97.7%;  Pred. No. 4.1e-53;
  Matches  251;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 ATGTGGGCGCTTTGCTCGCTGCTGCGGTCCGCGGCCGGACGCACCATGTCGCAGGGACGC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51 ATGTGGGCGCTTTGCTCGCTGCTGCGGTCCGCGGCCGGACGCACCATGTCGCAGGGACGC 110

Qy      61 ACCATATCGCAGGCACCCGCCCGCCGCGAGCGGCCGCGCAAGGACCCGCTGCGGCACCTG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     111 ACCATATCGCAGGCACCCGCCCGCCGCGAGCGGCCGCGCAAGGACCCGCTGCGGCACCTG 170

Qy     121 CGCACGCGAGAGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAAACACCGTGTACCTG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     171 CGCACGCGAGAGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAAACACCGTGTACCTG 230

Qy     181 CAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     231 CAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTC 290

Qy     241 AACCGGTATCTCTTCAA 257
           |||||||   ||  | |
Db     291 AACCGGTCAGTCAACGA 307


RESULT 15
AAA60207
ID   AAA60207 standard; DNA; 26664 BP.
XX
AC   AAA60207;
XX
DT   07-DEC-2000  (first entry)
XX
DE   Human prostate cancer predisposing gene HPC2 genomic sequence.
XX
KW   Human; prostate cancer predisposing gene; HPC2; chromosome 17p;
KW   gene therapy; peptide therapy; drug design; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             910..26039
FT                   /*tag= a
FT                   /product= "HPC2"
FT                   /note= "this sequence contains introns"
FT                   /transl_except= (pos:23892..23895,aa:Glu)
FT   exon            910..1154
FT                   /*tag= b
FT                   /number= 1
FT   exon            1736..1786
FT                   /*tag= c
FT                   /number= 2
FT   exon            1925..1995
FT                   /*tag= d
FT                   /number= 3
FT   exon            3025..3089
FT                   /*tag= e
FT                   /number= 4
FT   exon            4361..4418
FT                   /*tag= f
FT                   /number= 5
FT   exon            5582..5650
FT                   /*tag= g
FT                   /number= 6
FT   exon            7075..7194
FT                   /*tag= h
FT                   /number= 7
FT   exon            8186..8244
FT                   /*tag= i
FT                   /number= 8
FT   exon            12878..12936
FT                   /*tag= j
FT                   /number= 9
FT   exon            13032..13104
FT                   /*tag= k
FT                   /number= 10
FT   exon            13756..13868
FT                   /*tag= l
FT                   /number= 11
FT   exon            15283..15378
FT                   /*tag= m
FT                   /number= 12
FT   exon            16278..16416
FT                   /*tag= n
FT                   /number= 13
FT   exon            16498..16583
FT                   /*tag= o
FT                   /number= 14
FT   exon            18583..18701
FT                   /*tag= p
FT                   /number= 15
FT   exon            20349..20445
FT                   /*tag= q
FT                   /number= 16
FT   exon            22172..22310
FT                   /*tag= r
FT                   /number= 17
FT   exon            22879..22917
FT                   /*tag= s
FT                   /number= 18
FT   exon            23045..23154
FT                   /*tag= t
FT                   /number= 19
FT   exon            23795..23895
FT                   /*tag= u
FT                   /number= 20
FT   exon            23973..24093
FT                   /*tag= v
FT                   /number= 21
FT   exon            24354..24432
FT                   /*tag= w
FT                   /number= 22
FT   exon            25026..25170
FT                   /*tag= x
FT                   /number= 23
FT   exon            25812..26036
FT                   /*tag= y
FT                   /number= 24
FT   polyA_signal    26447..26452
FT                   /*tag= z
XX
PN   WO200027864-A1.
XX
PD   18-MAY-2000.
XX
PF   05-NOV-1999;   99WO-US26055.
XX
PR   06-NOV-1998;   98US-0107468.
```

XX
PA (MYRI-) MYRIAD GENETICS INC.
XX
PI Tavtigian SV, Teng DHF, Simard J, Rommens JM;
XX
DR WPI; 2000-376481/32.
DR P-PSDB; AAB07228.
XX
PT Human prostate cancer (HPC)2 nucleic acids, polypeptides, and
PT antibodies, useful for treatment and diagnosis of prostate cancer -
XX
PS Claim 3; Page 108-122; 157pp; English.
XX
CC The present sequence is the genomic sequence of the human prostate
CC cancer predisposing gene HPC2, which is found on chromosome 17p. Some
CC alleles of this gene cause a predisposition to cancer, particularly
CC prostate cancer. This gene and its protein can be used in peptide and
CC gene therapy for cancer patients, as well as being useful as diagnostic
CC tools (both for cancer sufferers and those with a predisposition to the
CC disease) and in the production of cancer drugs. This sequence was
CC isolated by cloning and sequencing the region of the genome which
CC appeared to cause a predisposition to prostate cancer.
XX
SQ Sequence 26664 BP; 6173 A; 6300 C; 6519 G; 7661 T; 11 other;

Query Match 10.0%; Score 247.4; DB 21; Length 26664;
Best Local Similarity 97.7%; Pred. No. 2.7e-52;
Matches 251; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy       1 ATGTGGGCGCTTTGCTCGCTGCTGCGGTCCGCGGCCGGACGCACCATGTCGCAGGGACGC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     910 ATGTGGGCGCTTTGCTCGCTGCTGCGGTCCGCGGCCGGACGCACCATGTCGCAGGGACGC 969

Qy      61 ACCATATCGCAGGCACCCGCCCGCCGCGAGCGGCCGCGCAAGGACCCGCTGCGGCACCTG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     970 ACCATATCGCAGGCACCCGCCCGCCGCGAGCGGCCGCGCAAGGACCCGCTGCGGCACCTG 1029

Qy     121 CGCACGCGAGAGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAAACACCGTGTACCTG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1030 CGCACGCGAGAGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAAACACCGTGTACCTG 1089

Qy     181 CAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1090 CAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTC 1149

Qy     241 AACCGGTATCTCTTCAA 257
           |||||||   ||  | |
Db    1150 AACCGGTCAGTCAACGA 1166
```

Search completed: May 16, 2003, 03:22:12
Job time : 574.621 secs

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: May 14, 2003, 09:56:52 ; Search time 18 Seconds
(without alignments)
1350.185 Million cell updates/sec

Title: US-09-434-382-2
Perfect score: 4325
Sequence: 1 MWALCSLLRSAAGRTMSQGR.....EPQQKRAHTEEPQAKKVRAQ 826

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4325	100.0	826	4	US-09-564-805-2	Sequence 2, Appli
2	4283	99.0	826	4	US-09-564-805-224	Sequence 224, App
3	4261	98.5	826	4	US-09-564-805-226	Sequence 226, App
4	3473.5	80.3	822	4	US-09-564-805-222	Sequence 222, App
5	875.5	20.2	837	4	US-09-564-805-228	Sequence 228, App
6	760	17.6	844	4	US-09-564-805-227	Sequence 227, App
7	599.5	13.9	838	4	US-09-315-794-52	Sequence 52, Appl
8	599.5	13.9	838	4	US-09-389-341-52	Sequence 52, Appl
9	599.5	13.9	838	4	US-09-564-805-229	Sequence 229, App
10	420	9.7	81	4	US-09-564-805-211	Sequence 211, App
11	281	6.5	307	4	US-09-564-805-232	Sequence 232, App
12	278	6.4	73	4	US-09-564-805-213	Sequence 213, App
13	275	6.4	311	4	US-09-564-805-230	Sequence 230, App
14	245.5	5.7	363	4	US-09-564-805-220	Sequence 220, App
15	243.5	5.6	326	4	US-09-564-805-231	Sequence 231, App
16	241.5	5.6	307	4	US-09-134-001C-3238	Sequence 3238, Ap
17	120.5	2.8	167	4	US-08-858-207A-353	Sequence 353, App
18	112	2.6	1093	5	PCT-US93-03077-1	Sequence 1, Appli
19	109.5	2.5	1649	4	US-09-535-008-75	Sequence 75, Appl
20	109.5	2.5	1650	4	US-09-535-008-71	Sequence 71, Appl
21	108.5	2.5	733	3	US-08-725-459B-28	Sequence 28, Appl
22	108	2.5	769	3	US-08-725-459B-39	Sequence 39, Appl
23	108	2.5	1141	1	US-08-363-300-2	Sequence 2, Appli
24	107	2.5	556	4	US-09-011-762-7	Sequence 7, Appli
25	106.5	2.5	733	3	US-08-725-459B-30	Sequence 30, Appl
26	106.5	2.5	2205	1	US-08-093-453B-2	Sequence 2, Appli
27	105	2.4	630	3	US-08-725-459B-17	Sequence 17, Appl
28	105	2.4	680	3	US-08-725-459B-3	Sequence 3, Appli
29	105	2.4	730	3	US-08-725-459B-2	Sequence 2, Appli
30	105	2.4	733	3	US-08-725-459B-29	Sequence 29, Appl
31	105	2.4	769	3	US-08-725-459B-35	Sequence 35, Appl
32	105	2.4	773	1	US-08-524-757-6	Sequence 6, Appli
33	105	2.4	773	3	US-08-725-459B-1	Sequence 1, Appli
34	105	2.4	773	3	US-08-725-459B-79	Sequence 79, Appl
35	103	2.4	548	2	US-08-452-075-3	Sequence 3, Appli
36	103	2.4	548	4	US-09-231-061-2	Sequence 2, Appli
37	102	2.4	769	3	US-08-725-459B-40	Sequence 40, Appl
38	102	2.4	1356	4	US-09-098-707A-2	Sequence 2, Appli
39	102	2.4	1356	4	US-09-483-539-2	Sequence 2, Appli
40	102	2.4	2544	2	US-08-576-626A-32	Sequence 32, Appl
41	100	2.3	769	3	US-08-725-459B-38	Sequence 38, Appl
42	99.5	2.3	1142	2	US-08-993-118-7	Sequence 7, Appli
43	99.5	2.3	1142	3	US-08-845-528C-7	Sequence 7, Appli
44	99.5	2.3	1142	4	US-09-061-709-2	Sequence 2, Appli
45	99	2.3	748	3	US-08-725-459B-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1
US-09-564-805-2
; Sequence 2, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-564-805-2

Query Match 100.0%; Score 4325; DB 4; Length 826;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 826; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MWALCSLLRSAAGRTMSQGRTISQAPARRERPRKDPLRHLRTREKRGPSGCSGGPNTVYL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MWALCSLLRSAAGRTMSQGRTISQAPARRERPRKDPLRHLRTREKRGPSGCSGGPNTVYL 60

Qy      61 QVVAAGSRDSGAALYVFSEFNRYLFNCGEGVQRLMQEHKLKVARLDNIFLTRMHWSNVGG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 QVVAAGSRDSGAALYVFSEFNRYLFNCGEGVQRLMQEHKLKVARLDNIFLTRMHWSNVGG 120

Qy     121 LSGMILTLKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPEYEDETMTV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 LSGMILTLKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPEYEDETMTV 180

Qy     181 YQIPIHSEQRRGKHQPWQSPERPLSRLSPERSSDSESNENEPHLPHGVSQRRGVRDSSLV 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 YQIPIHSEQRRGKHQPWQSPERPLSRLSPERSSDSESNENEPHLPHGVSQRRGVRDSSLV 240

Qy     241 VAFICKLHLKRGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSITHEGREILAEELCTPP 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 VAFICKLHLKRGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSITHEGREILAEELCTPP 300
```

```
Qy    301 DPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPASVLVDSRYQQWMER 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 DPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPASVLVDSRYQQWMER 360

Qy    361 FGPDTQHLVLNENCASVHNLRSHKIQTQLNLIHPDIFPLLTSFRCKKEGPTLSVPMVQGE 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 FGPDTQHLVLNENCASVHNLRSHKIQTQLNLIHPDIFPLLTSFRCKKEGPTLSVPMVQGE 420

Qy    421 CLLKYQLRPRREWQRDAIITCNPEEFIVEALQLPNFQQSVQEYRRSAQDGPAPAEKRSQY 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CLLKYQLRPRREWQRDAIITCNPEEFIVEALQLPNFQQSVQEYRRSAQDGPAPAEKRSQY 480

Qy    481 PEIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDCGEGTFGQLCRHYGDQVDRVLGTLA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 PEIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDCGEGTFGQLCRHYGDQVDRVLGTLA 540

Qy    541 AVFVSHLHADHHTGLPSILLQRERALASLGKPLHPLLVVAPNQLKAWLQQYHNQCQEVLH 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AVFVSHLHADHHTGLPSILLQRERALASLGKPLHPLLVVAPNQLKAWLQQYHNQCQEVLH 600

Qy    601 HISMIPAKCLQEGAEISSPAVERLISSLLRTCDLEEFQTCLVRHCKHAFGCALVHTSGWK 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 HISMIPAKCLQEGAEISSPAVERLISSLLRTCDLEEFQTCLVRHCKHAFGCALVHTSGWK 660

Qy    661 VVYSGDTMPCEALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMRMNAEFI 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 VVYSGDTMPCEALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMRMNAEFI 720

Qy    721 MLNHFSQRYAKVPLFSPNFSEKVGVAFDHMKVCFGDFPTMPKLIPPLKALFAGDIEEMEE 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 MLNHFSQRYAKVPLFSPNFSEKVGVAFDHMKVCFGDFPTMPKLIPPLKALFAGDIEEMEE 780

Qy    781 RREKRELRQVRAALLSRELAGGLEDGEPQQKRAHTEEPQAKKVRAQ 826
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    781 RREKRELRQVRAALLSRELAGGLEDGEPQQKRAHTEEPQAKKVRAQ 826
```

RESULT 2
US-09-564-805-224
; Sequence 224, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 224
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Pan troglodytes
US-09-564-805-224

Query Match 99.0%; Score 4283; DB 4; Length 826;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 817; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```
Qy      1 MWALCSLLRSAAGRTMSQGRTISQAPARRERPRKDPLRHLRTREKRGPSGCSGGPNTVYL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MWALCSLLRSAAGRTMSQGRTISQAPARRERPRKDPLRHLRTREKRGPSGCSGGPNTVYL 60

Qy     61 QVVAAGSRDSGAALYVFSEFNRYLFNCGEGVQRLMQEHKLKVARLDNIFLTRMHWSNVGG 120
          ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db     61 QVVAAGSRDSGAALYVFSEFNRYLFNCGEGIQRLMQEHKLKVARLDNIFLTRMHWSNVGG 120

Qy    121 LSGMILTLKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPEYEDETMTV 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 LSGMILTLKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPEYEDETMTV 180

Qy    181 YQIPIHSEQRRGKHQPWQSPERPLSRLSPERSSDSESNENEPHLPHGVSQRRGVRDSSLV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 YQIPIHSEQRRGKHQPWQSPERPLSRLSPERSSDSESNENEPHLPHGVSQRRGVRDSSLV 240

Qy    241 VAFICKLHLKRGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSITHEGREILAEELCTPP 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 VAFICKLHLKRGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSITHEGREILAEELCTPP 300

Qy    301 DPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPESVLVDSRYQQWMER 360
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    301 DPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPESVLVDSRYQQWMER 360

Qy    361 FGPDTQHLVLNENCASVHNLRSHKIQTQLNLIHPDIFPLLTSFRCKKEGPTLSVPMVQGE 420
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    361 FGPDTQHLVLNENCASVHNLRSHKIQTQLNLIHPDIFPLLTSFPCKKEGPTLSVPMVQGE 420

Qy    421 CLLKYQLRPRREWQRDAIITCNPEEFIVEALQLPNFQQSVQEYRRSAQDGPAPAEKRSQY 480
          ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db    421 CLLKYQLRPRREWQRDAIITCNPEEFIIEALQLPNFQQSVQEYRRSAQDGPAPAEKRSQY 480

Qy    481 PEIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDCGEGTFGQLCRHYGDQVDRVLGTLA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 PEIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDCGEGTFGQLCRHYGDQVDRVLGTLA 540

Qy    541 AVFVSHLHADHHTGLPSILLQRERALASLGKPLHPLLVVAPNQLKAWLQQYHNQCQEVLH 600
          |||||||||||||| :||||||||||||| ||||||||||||||||||||||||||||||
Db    541 AVFVSHLHADHHTGLLNILLQRERALASLGKPFHPLLVVAPNQLKAWLQQYHNQCQEVLH 600

Qy    601 HISMIPAKCLQEGAEISSPAVERLISSLLRTCDLEEFQTCLVRHCKHAFGCALVHTSGWK 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 HISMIPAKCLQEGAEISSPAVERLISSLLRTCDLEEFQTCLVRHCKHAFGCALVHTSGWK 660

Qy    661 VVYSGDTMPCEALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMRMNAEFI 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 VVYSGDTMPCEALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMRMNAEFI 720

Qy    721 MLNHFSQRYAKVPLFSPNFSEKVGVAFDHMKVCFGDFPTMPKLIPPLKALFAGDIEEMEE 780
          ||||||||||||||||:|||||||||||||| ||||||||||||||||||||||||||||||
Db    721 MLNHFSQRYAKVPLFSPNFNEKVGVAFDHMKVCFGDFATMPKLIPPLKALFAGDIEEMEE 780

Qy    781 RREKRELRQVRAALLSRELAGGLEDGEPQQKRAHTEEPQAKKVRAQ 826
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    781 RREKRELRQVRAALLSRELAGGLEDGEPQQKRAHTEEPQAKKVRAQ 826
```

RESULT 3
US-09-564-805-226
; Sequence 226, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240

```
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 226
;   LENGTH: 826
;   TYPE: PRT
;   ORGANISM: Gorilla gorilla
US-09-564-805-226

  Query Match             98.5%;  Score 4261;  DB 4;  Length 826;
  Best Local Similarity   98.5%;  Pred. No. 0;
  Matches  814;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      1 MWALCSLLRSAAGRTMSQGRTISQAPARRERPRKDPLRHLRTREKRGPSGCSGGPNTVYL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MWALCSLLRSAAGRTMSQGRTISQAPARRERPRKDPLRHLRTREKRGPSGCSGGPNTVYL 60

Qy     61 QVVAAGSRDSGAALYVFSEFNRYLFNCGEGVQRLMQEHKLKVARLDNIFLTRMHWSNVGG 120
          ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     61 QVVAAGSRDSGAALYVFSEFNRYLFNCGEGVQRLMQEHKLKVVRLDNIFLTRMHWSNVGG 120

Qy    121 LSGMILTLKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPEYEDETMTV 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 LSGMILTLKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPEYEDETMTV 180

Qy    181 YQIPIHSEQRRGKHQPWQSPERPLSRLSPERSSDSESNENEPHLPHGVSQRRGVRDSSLV 240
          ||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 YQIPIHSEQRRGRHQPWQSPERPLSRLSPERSSDSESNENEPHLPHGVSQRRGVRDSSLV 240

Qy    241 VAFICKLHLKRGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSITHEGREILAEELCTPP 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 VAFICKLHLKRGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSITHEGREILAEELCTPP 300

Qy    301 DPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPASVLVDSRYQQWMER 360
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    301 DPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPESVLVDSRYQQWMER 360

Qy    361 FGPDTQHLVLNENCASVHNLRSHKIQTQLNLIHPDIFPLLTSFRCKKEGPTLSVPMVQGE 420
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    361 FGPDTQHLVLNENCASVHNLRSHKIQTQLNLIHPDIFPLLTSFPCKKEGPTLSVPMVQGE 420

Qy    421 CLLKYQLRPRREWQRDAIITCNPEEFIVEALQLPNFQQSVQEYRRSAQDGPAPAEKRSQY 480
          |||||||||||||||||||||||||||||||||||||||||| || ||||||||||||||||
Db    421 CLLKYQLRPRREWQRDAIITCNPEEFIVEALQLPNFQQSVQEYRRSVQDVPAPAEKRSQY 480

Qy    481 PEIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDCGEGTFGQLCRHYGDQVDRVLGTLA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 PEIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDCGEGTFGQLCRHYGDQVDRVLGTLA 540

Qy    541 AVFVSHLHADHHTGLPSILLQRERALASLGKPLHPLLVVAPNQLKAWLQQYHNQCQEVLH 600
          |||||||||||||||:||||||:||||||||||||||||:||||||||||||||||||||||
Db    541 AVFVSHLHADHHTGLLNILLQREQALASLGKPLHPLLVVAPSQLKAWLQQYHNQCQEVLH 600

Qy    601 HISMIPAKCLQEGAEISSPAVERLISSLLRTCDLEEFQTCLVRHCKHAFGCALVHTSGWK 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 HISMIPAKCLQEGAEISSPAVERLISSLLRTCDLEEFQTCLVRHCKHAFGCALVHTSGWK 660

Qy    661 VVYSGDTMPCEALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMRMNAEFI 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 VVYSGDTMPCEALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMRMNAEFI 720

Qy    721 MLNHFSQRYAKVPLFSPNFSEKVGVAFDHMKVCFGDFPTMPKLIPPLKALFAGDIEEMEE 780
          ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db    721 MLNHFSQRYAKVPLFSPNFNEKVGVAFDHMKVCFGDFPTMPKLIPPLKALFAGDIEEMEE 780

Qy    781 RREKRELRQVRAALLSRELAGGLEDGEPQQKRAHTEEPQAKKVRAQ 826
          |||||||||||||| ||||||||||||||||||||||||||||||||
Db    781 RREKRELRQVRAALLSGELAGGLEDGEPQQKRAHTEEPQAKKVRAQ 826


RESULT 4
US-09-564-805-222
; Sequence 222, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
;  APPLICANT:  Tavtigian, Sean V.
;  APPLICANT:  Teng, David H.F.
;  APPLICANT:  Simard, Jacques
;  APPLICANT:  Rommens, Johanna M.
;  APPLICANT:  Myriad Genetics, Inc.
;  TITLE OF INVENTION:  Chromosome 17p-Linked Prostate Cancer Susceptibility
;  TITLE OF INVENTION:  Gene and a Paralog and Orthologous Genes
;  FILE REFERENCE:  2318-258
;  CURRENT APPLICATION NUMBER: US/09/564,805
;  CURRENT FILING DATE:  2000-05-05
;  PRIOR APPLICATION NUMBER:  US 60/107,468
;  PRIOR FILING DATE:  1998-11-06
;  PRIOR APPLICATION NUMBER:  09/434,382
;  PRIOR FILING DATE:  1999-11-05
;  NUMBER OF SEQ ID NOS:  240
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 222
;   LENGTH: 822
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-564-805-222

  Query Match             80.3%;  Score 3473.5;  DB 4;  Length 822;
  Best Local Similarity   80.5%;  Pred. No. 0;
  Matches  665;  Conservative   66;  Mismatches   76;  Indels   19;  Gaps    6;

Qy      1 MWALCSLLRSAAGRTMSQGRTISQAPARRERPRKDPLRHLRTREKRGPSGCSGGPNTVYL 60
          |||| ||||    ||||||        ||| || ||||||||||||||    ||||||||
Db      1 MWALRSLLRPLGLRTMSQG------SARRPRPPKDPLRHLRTREKRGPG--PGGPNTVYL 52

Qy     61 QVVAAGSRDSGAALYVFSEFNRYLFNCGEGVQRLMQEHKLKVARLDNIFLTRMHWSNVGG 120
          ||||||  ||:|||||||||:|||||||||||||||||||  :   ||||||||||||||
Db     53 QVVAAGGRDAGAALYVFSEYNRYLFNCGEGVQRLMQEHKTESRSLDNIFLTRMHWSNVGG 112

Qy    121 LSGMILTLKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPEYKDETMTV 180
          | ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db    113 LCGMILTLKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPEYKDETMTV 172

Qy    181 YQIPIHSEQRRGKHQPWQSPERPLSRLSPERSSDSESNENEPHLPHGVSQRRGV-RDSSL 239
          ||:|||||:| || || |||    :||||::|||| | ||      |  |:  : :  ||
Db    173 YQVPIHSERRCGKQQPSQSPRTSPNRLSPKQSSDSGSAEN------GQCQQESMGQGPSL 226

Qy    240 VVAFICKLHLKRGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSITHEGREILAEELCTP 299
          ||||:|||||::||||||||||:|||||||||||||||||||||:||||| |||||||
Db    227 VVAFVCKLHLRKGNFLVLKAKELGLPVGTAAIAPIIAAVKDGKSITYEGREIAAEELCTP 286

Qy    300 PDPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPASVLVDSRYQQWME 359
          ||||  |:|||||| || ||||| ||:||| :||||||||||:|| |||:|||||||||
Db    287 PDPGLVFIVVECPDEGFILPICENDTFKRYQAEADAPVALVVHIAPESVLIDSRYQQWME 346

Qy    360 RFGPDTQHLVLNENCASVHNLRSHKIQTQLNLIHPDIFPLLTSFRCKKEGPTLSVPMVQG 419
          |||||||||:|||| ||||||||||||||:|||||||| ||||  |:|| |||| |:|
Db    347 RFGPDTQHLILNENCPSVHNLRSHKIQTQLSLIHPDIFPQLTSFYSKEEGSTLSVPTVRG 406

Qy    420 ECLLKYQLRPRREWQRDAIITCNPEEFIVEALQLPNFQQSVQEYRRSAQDGPAPAEKRSQ 479
          |||||| :||:|||||  : || :||| |||:||:||:||:|||:: |: |||||||||
Db    407 ECLLKYSVRPKREWQRDTTLDCNTDEFIAEALELPSFQESVEEYRKNVQENPAPAEKRSQ 466

Qy    480 YPEIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDCGEGTFGQLCRHYGDQVDRVLGTL 539
          ||||:|||||||||:|||||:||||:||| |:|||||||||||||| |:|||| :|
Db    467 YPEIVFLGTGSAIPMEIRNVSSTLVNLSPDKSVLLDCGEGTFGQLCRHYGQQIDRVLCSL 526

Qy    540 AAVFVSHLHADHHTGLPSILLQRERALASLGKPLHPLLVVAPNQLKAWLQQYHNQCQEVL 599
           |||||||||||||| :|||||| ||||||| |||||| ||:|||||| |||:|
Db    527 TAVFVSHLHADHHTGLLNILLQREHALASLGKPFQPLLVVAPTQLRAWLQQYHNHCQEIL 586

Qy    600 HHISMIPAKCLQEGAEISSPAVERLISSLLRTCDLEEFQTCLVRHCKHAFGCALVHTSGW 659
          ||:|||||||||:|||:|:  :||||| || ||||||||||||||||||||||||||:|||
Db    587 HHVSMIPAKCLQKGAEVSNTTLERLISLLLETCDLEEFQTCLVRHCKHAFGCALVHSSGW 646

Qy    660 KVVYSGDTMPCEALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMRMNAEF 719
          |||||||||||||:|||||||||||||| |||||||:|||||||:|||||||
Db    647 KVVYSGDTMPCEALVQMGKDATLLIHEATLEDXLEEEAVERTHSTTSQAINVGMRMNAEF 706
```

```
Qy    720 IMLNHFSQRYAKVPLFSPNFSEKVGVAFDHMKVCFGDFPTMPKLIPPLKALFAGDIEEME 779
          |||||||||| |:|||||:|:||||:||||||| |||||||:|||||||||||||||||||
Db    707 IMLNHFSQRYXKIPLFSPDFNEKVGIAFDHMKVXFGDFPTVPKLIPPLKALFAGDIEEMV 766

Qy    780 ERREKRELRQVRAALLSRELAGGLEDGEPQQKRAHTEE---PQAKK 822
          ||||||||| ||||||::: |   || ||||||||||:|   ||:||
Db    767 ERREKRELRLVRAALLTQQ-ADSPEDREPQQKRAHTDEPHSPQSKK 811


RESULT 5
US-09-564-805-228
; Sequence 228, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
;  APPLICANT:  Tavtigian, Sean V.
;  APPLICANT:  Teng, David H.F.
;  APPLICANT:  Simard, Jacques
;  APPLICANT:  Rommens, Johanna M.
;  APPLICANT:  Myriad Genetics, Inc.
;  TITLE OF INVENTION:  Chromosome 17p-Linked Prostate Cancer Susceptibility
;  TITLE OF INVENTION:  Gene and a Paralog and Orthologous Genes
;  FILE REFERENCE:  2318-258
;  CURRENT APPLICATION NUMBER: US/09/564,805
;  CURRENT FILING DATE:  2000-05-05
;  PRIOR APPLICATION NUMBER:  US 60/107,468
;  PRIOR FILING DATE:  1998-11-06
;  PRIOR APPLICATION NUMBER:  09/434,382
;  PRIOR FILING DATE:  1999-11-05
;  NUMBER OF SEQ ID NOS:  240
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 228
;   LENGTH: 837
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-09-564-805-228

  Query Match             20.2%;  Score 875.5;  DB 4;  Length 837;
  Best Local Similarity   29.4%;  Pred. No. 1.8e-76;
  Matches  250;  Conservative  128;  Mismatches  297;  Indels  175;  Gaps   28;

Qy     41 RTREKRGPSGCSGGPNTV-YLQVVAAG--SRDSGAALYVFSEFNRYLFNCGEGVQRLMQE 97
          |  :|  |:      ||: | |::  |  ::|: ::: :| :  |::|| |||:||   |
Db     39 RKSQKLNPT------NTIAYAQILGTGMDTQDTSSSVLLFFDKQRFIFNAGEGLQRFCTE 92

Qy     98 HKLKVARLDNIFLTRMHWSNVGGLSGMILTLK---ETGLPKCVLSGPPQLEKYLEAIKIF 154
          ||:|::::|::||:|:     ||| |::|||    | ||   |  ||  |   ::|:| | .
Db     93 HKIKLSKIDHVFLSRVCSETAGGLPGLLLTLAGIGEEGLSVNVW-GPSDLNYLVDAMKSF 151

Qy    155 SGPLKGIEL-AVRPHSAPE----YEDETMTVYQI---PIHSEQRRGKHQPWQSPERPLSR 206
             :   :  | | |:      || : :  |   | |||:
Db    152 IPRAAMVHTRSFGPSSTPDPIVLVNDEVVKISAIILKPCHSEE----------------- 194

Qy    207 LSPERSSDSESNENEPHLPHGVSQRRGVRDSSLVVAFICKLHLKRGNFLVLKAKEM-GLP 265
             ||                | :   | | ::|:|    | | : |||:: |:
Db    195 -------DS----------------GNKSGDLSVVYVCELPEILGKFDLEKAKKVFGVK 230

Qy    266 VGTAAIAPIIAAVKDGKSITHEGREILA--EELCTPPDPGAAFVVVECPDESFIQPICEN 323
          |     |  : :: |:|:  : |:|     ::  |  ||   ::|:|| ||    :
Db    231 PG-----PKYSRLQSGESVKSDERDITVHPSDVMGPSLPGPIVLLVDCPTESHAAELFSL 285

Qy    324 ATFQRYQGKADAP------VALVVHMAPASVLVDSRYQQWMERFGPDTQHLV-------- 369
          : : |    |         |  ::|::|:||    || ||::|   |||::
Db    286 KSLESYYSSPDEQTIGAKFVNCIIHLSPSSVTSSPTYQSWMKKFHL-TQHILAGHQRFLP 344

Qy    370 ----------LNENCASVHNLRSHKIQTQLNLIHPDIFPLLTSFRCKKEGPTLSVPMVQG 419
                   : :| |     | :|  :|| : |  ||    :  :    ::  |
Db    345 LLIIVSHQKTVRKNMAFPILKASSRIAARLNYLCPQFFPAPGFWPSQLTDNSIIDPTPSN 404

Qy    420 ECLLKYQLRP--RREWQRDAIITCNPEEFIVEAL--QLPNFQQSVQEYRR--SAQDGPAP 473
            |: |||   |   |  |        :|: |  ::|  :   :| ::  : |
Db    405 ----KFNLRPVAIRGIDRSCIPAPLTSSEVVDELLSEIPEIKDKSEEIKQFWNKQHNKTI 460
```

```
Qy    474 AEK----------------RSQYPEIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDC 516
          ||                |    ||: |||||: | | ||||  :::    ||||||
Db    461 IEKLWLSECNTVLPNCLEKIRRDDMEIVILGTGSSQPSKYRNVSAIFIDLFSRGSLLLDC 520

Qy    517 GEGTFGQLCRHYG-DQVDRVLGTLAAVFVSHLHADHHTGLPSILLQRERALASLGKPLHP 575
          |||| ||| | || |  |  :  |  :::||:|||||||  ||  | : |   |    |
Db    521 GEGTLGQLKRRYGLDGADEAVRKLRCIWISHIHADHHTGLARILALRSKLLK--GVTHEP 578

Qy    576 LLVVAPNQLKAWLQQYHNQCQEVLHHISMIPAKC----------LQEGAEI-------SS 618
          ::|| |  || :|  |       |  : |    |           |: | |         |
Db    579 VIVVGPRPLKRFLDAYQR-----LEDLDMEFLDCRSTTATSWASLESGGEAEGSLFTQGS 633

Qy    619 PAVE-----------------RLISSLLRTCDLEEFQTCLVRHCKHAFGCALVHTS--- 657
          |      .              : :  :|    | :  :  | ||  |:|  :
Db    634 PMQSVFKRSDISMDNSSVLLCLKNLKKVLSEIGLNDLISFPVVHCPQAYGVVIKAAERVN 693

Qy    658 -------GWKVVYSGDTMPCEALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAIS 710
                |||:|||||: ||    |   :|||:|||||| || | |||: | |||| :||
Db    694 SVGEQILGWKMVYSGDSRPCPETVEASRDATILIHEATFEDALIEEALAKNHSTTKEAID 753

Qy    711 VGMRMNAEFIMLNHFSQRYAKVPLFSPNFSEKVGVAFDHMKVCFGDFPTMPKLIPPLKAL 770
          ||   |   |:| ||||||  |:|:   :       :||| | :   |   :||::|  | |
Db    754 VGSAANVYRIVLTHFSQRYPKIPVIDESHMHNTCIAFDLMSINMADLHVLPKVLPYFKTL 813

Qy    771 FAGDIEEMEE 780
          |  :: | |:
Db    814 FRDEMVEDED 823


RESULT 6
US-09-564-805-227
; Sequence 227, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
;  APPLICANT:  Tavtigian, Sean V.
;  APPLICANT:  Teng, David H.F.
;  APPLICANT:  Simard, Jacques
;  APPLICANT:  Rommens, Johanna M.
;  APPLICANT:  Myriad Genetics, Inc.
;  TITLE OF INVENTION:  Chromosome 17p-Linked Prostate Cancer Susceptibility
;  TITLE OF INVENTION:  Gene and a Paralog and Orthologous Genes
;  FILE REFERENCE:  2318-258
;  CURRENT APPLICATION NUMBER: US/09/564,805
;  CURRENT FILING DATE:  2000-05-05
;  PRIOR APPLICATION NUMBER:  US 60/107,468
;  PRIOR FILING DATE:  1998-11-06
;  PRIOR APPLICATION NUMBER:  09/434,382
;  PRIOR FILING DATE:  1999-11-05
;  NUMBER OF SEQ ID NOS:  240
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 227
;   LENGTH: 844
;   TYPE: PRT
;   ORGANISM: Caenorhabditis elegans
US-09-564-805-227

  Query Match             17.6%;  Score 760;  DB 4;  Length 844;
  Best Local Similarity   26.6%;  Pred. No. 3.8e-65;
  Matches  226;  Conservative  175;  Mismatches  324;  Indels  126;  Gaps   29;

Qy      9 RSAAGRTMSQGRTISQAPARRERPRKDPLRH----LRTREKRGPSGCSGG----PNTVYL 60
          | |  | : |  : |   ||         ||     :: ::|      :      |: | :
Db     36 RIARNRRILQKHSSSHLKAREVNASISNLRQSMAAVQKKQKAAHEPPANSIVNIPSQVSI 95

Qy     61 QVVAAGSRDSGAALYVFSEFNRYLFNCGEGVQRLMQEHKLKVARLDNIFLTRMHWSNVGG 120
          :|:  |:    |   : :    |:||| |   | : : ::: : : ::|:|  :| |: |
Db     96 EVLGNGTGLLRACFILRTPLKTYMFNCPENACRFLWQLRIRSSSVVDLFITSANWDNIAG 155

Qy    121 LSGMILTLKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKG-----IELAVRPHSAPEYED 175
          :| ::|: ||:      | |   :: :|| |: |      |      ::  ||::   |||
Db    156 ISSILLS-KESNALSTRLHGAMNIKHFLECIRPFQDSDYGSCKYPSQVEERPYTMENYED 214

Qy    176 ETMTVYQIPIHSEQRRGKHQPWQSPERPLSRLSPERSSDSESNENEPHLPHGVSQRRGVR 235
```

```
Db    215 AGLKVTYIPL-------------SP--PLN-----IGSNNEKSKN-------------VK 241

Qy    236 DSSLVVAFICKLHLKRGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSIT-HEGREILAE 294
Db    242 VNNVDIAFLIEMKEAARRIDTMKLMELKVPKG-----PLIGKLKSGEAVTLPDGRTIQPD 296

Qy    295 ELCTP---PDPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPASVLVD 351
Db    297 QVFSSDKVEGDKPLLLVTECTTEDHVKALIDSSSLQPFL-NGEKQLDYMVHISDDAVINT 355

Qy    352 SRYQQWMERF-GPDTQHLVLNENCASVHNLRS-HKIQTQLNLIHPDIFPLLTSFRCKKEG 409
Db    356 PTYRHLMEKLNNPSITHLLINGGNPVIPAVESVYKHTRLLRSIAPSLFPALHPI------ 409

Qy    410 PTLSVPMVQGECLLKYQ-----LRP-RREWQRDAIITCNPEEFIVEALQL-----PNFQQ 458
Db    410 -DWSGIITQNEELSQRQDQFIRVAPMQRYWMRRG-ASFNEEPIVNNLLAAEPELSDKAKE 467

Qy    459 SVQEYRRSAQDGPAPAEKRSQYPEIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDCGE 518
Db    468 LIKEYQKLEKENKMDCE----FPKLTFFGTSSAVPSKYRNVTGYLVEASENSAILIDVGE 523

Qy    519 GTFGQLCRHYG-DQVDRVLGTLAAVFVSHLHADHHTGLPSILLQRERALASLGKPLHPLL 577
Db    524 GTYGQMRAVFGEDGCKQLLVNLNCVLITHAHQDHMNGLYTIIARRKEAFESLGAPYRPLV 583

Qy    578 VVAPNQLKAWLQQYHNQCQEVLHHISMI-------------PAKCLQEGAEISSP----- 619
Db    584 LVCNRNVLKPMKTY-SICFENIEHLLEIVDISRYPLTPPGSPGGPPGKRPRLPSPHLPPS 642

Qy    620 --AVERLISSL-LRTCDLEEFQTCLVRHCKHAFGCALVHTSGWKVVYSGDTMPCEALVRM 676
Db    643 RDVLQDMSSSFDKKAWKLDELKAVQVHHTRMANG-FVMRVAGKRIVFSGDTKPCDLLVEE 701

Qy    677 GKDATLLIHEATLEDGLE----------------EEAVEKTHSTTSQAISVGMRMNAEF 719
Db    702 GKDADVLVHESTFEDGHEVDMTPKPPKKLAKISSLADAMRKRHSTMGQAVDVGKRMNAKH 761

Qy    720 IMLNHFSQRYAKVPLFSPNF--SEKVGVAFDHMKVCFGDFPTMPKLIPPLKALFAGDIEE 777
Db    762 IILTHFSARYPKVPVL-PEYLDKENIGVAMDMLRVRFDHLPLVSKLLPIFREVFVAELFE 820

Qy    778 MEERREKRELR 788
Db    821 LTIKKEQRVLK 831
```

RESULT 7
US-09-315-794-52
; Sequence 52, Application US/09315794
; Patent No. 6197517
; GENERAL INFORMATION:
; APPLICANT: Roberts, Christopher J.
; TITLE OF INVENTION: ESSENTIAL GENES OF YEAST AS TARGETS FOR ANTIFUNGAL
; TITLE OF INVENTION: AGENTS, HERBICIDES, INSECTICIDES AND ANTI-PROLIFERATION
; TITLE OF INVENTION: DRUGS
; FILE REFERENCE: 9301-053
; CURRENT APPLICATION NUMBER: US/09/315,794
; CURRENT FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
; LENGTH: 838
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-315-794-52

Query Match 13.9%; Score 599.5; DB 4; Length 838;
Best Local Similarity 25.7%; Pred. No. 1.9e-49;
Matches 221; Conservative 138; Mismatches 290; Indels 211; Gaps 36;

```
Qy     82 RYLF-NCGEGVQRLMQEHKLKVARLDNIFLT-RMHWSNVGGLSGMILTLKETGLPKCVLS 139

Db     28 KYFFGKIGEGSQRSLTENKIRISKLKDIFLTGELNWSDIGGLPGMILTIADQGKSNLVLH 87

Qy    140 GPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPE---YEDETMTVYQIPI---HSEQRRGK 193
Db     88 YGNDILNYIVSTWRYFVFRFGIDL--NDHIMKDKEVYKDKIIAVKSFNVLKNGGEDRLGV 145

Qy    194 HQPWQS------------PERPLSRLSPERSSDSESNENEPHLPHGVSQRRGVRDSSLVV 241
Db    146 FDSFQKGVLRSIVAKMFPKHAPTDRYDP--SSDPHLNVELPDL-----------DAKVEV 192

Qy    242 AFICKLHLK--RGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSIT-HEGREILAEELCT 298
Db    193 STNYEISFSPVRGKFKVEEAIKLGVPKG-----PLFAKLTKGQTITLDNGIVVTPEQVLE 247

Qy    299 PPDPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPASVLVDSRYQQWM 358
Db    248 NERHFAKVLILDIPDDLYL-----NAFVEKFKDYDCAELGMVYYFLGDEVTINDNLFAFI 302

Qy    359 ERFGPDTQHLVLNENCASVHNLRSHKIQTQLNLIHPDIFPL----LTSFRCK-------- 406
Db    303 DIFE--------KNNYGKVNHMISH------NKISPNTISFFGSALTTLKLKALQVNNYN 348

Qy    407 --KEGPTLS--------VPMVQGECLLKYQLRPRRE---------WQRDAIITCNP---- 443
Db    349 LPKTDRVFSKDFYDRFDTPLSRGTSMCKSQEEPLNTIIEKDNIHIFSQNKTVTFEPFRMN 408

Qy    444 ----------------------EEFIVEALQLP--------NFQQSVQEYRRSAQDGPAPA 474
Db    409 EEPMKCNINGEVADFSWQEIFEEH-VKPLEFPLADVDTVINNQLHVDNFNNSAE------ 461

Qy    475 EKRSQYPEIIFLGTGSAIPMKIRNVSATLVNI--------SPDTSLLLDCGEGTFGQLCR 526
Db    462 --KKKHVEIITLGTGSALPSKYRNVVSTLVKVPFTDADGNTINRNIMLDAGENTLGTIHR 519

Qy    527 HYGD-QVDRVLGTLAAVFVSHLHADHHTGLPSILLQRERALASLGKPLHPLLVVAPNQLK 585
Db    520 MFSQLAVKSIFQDLKMIYLSHLHADHHLGIISVL--NEWYKYNKDDETSYIYVVTP---- 573

Qy    586 AWLQQYHN--------QCQEVLHHISMIPA----------------------KCLQEGA- 614
Db    574 -W--QYHKFVNEWLVLENKEILKRIKYISCEHFINDSFVRMQTQSVPLAEFNEILKENSN 630

Qy    615 -------EISSPAVER---LISSLLRTCDLEEFQTCLVRHCKHAFGCALV--------HT 656
Db    631 QESNRKLELDRDSSYRDVDLIRQMYEDLSIEYFQTCRAIHCDWAYSNSITFRMDENNEHN 690

Qy    657 SGWKVVYSGDTMPC--EALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMR 714
Db    691 T-FKVSYSGDTRPNIEKFSLEIGYNSDLLIHEATLENQLLEDAVKKKHCTINEAIGVSNK 749

Qy    715 MNAEFIMLNHFSQRYAKVPLFSPN---FSEKVGVAFDHMKVCFGDFPTMPKLIPPLKALF 771
Db    750 MNARKLILTHFSQRYPKLPQLDNNIDVMAREFCFAFDSMIVDYEKIGEQQRIFPLLNKAF 809

Qy    772 AGDIEEMEERREKRELRQVR 791
Db    810 ---VEEKEEEEDVDDVESVQ 826
```

RESULT 8
US-09-389-341-52
; Sequence 52, Application US/09389341
; Patent No. 6200803
; GENERAL INFORMATION:
; APPLICANT: Roberts, Christopher J.
; TITLE OF INVENTION: ESSENTIAL GENES OF YEAST AS TARGETS FOR ANTIFUNGAL
; TITLE OF INVENTION: AGENTS, HERBICIDES, INSECTICIDES AND ANTI-PROLIFERATIVE
; TITLE OF INVENTION: DRUGS
; FILE REFERENCE: 9301-057
; CURRENT APPLICATION NUMBER: US/09/389,341
; CURRENT FILING DATE: 1999-09-02
; EARLIER APPLICATION NUMBER: 09/315,794

```
;  EARLIER FILING DATE: 1999-05-21
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
;   LENGTH: 838
;   TYPE: PRT
;   ORGANISM: Saccharomyces cerevisiae
US-09-389-341-52

  Query Match             13.9%;  Score 599.5;  DB 4;  Length 838;
  Best Local Similarity   25.7%;  Pred. No. 1.9e-49;
  Matches  221;  Conservative  138;  Mismatches  290;  Indels  211;  Gaps   36;

Qy      82 RYLF-NCGEGVQRLMQEHKLKVARLDNIFLT-RMHWSNVGGLSGMILTLKETGLPKCVLS 139
           :| |   ||| || : |:|:::::| :|||| ::||::||| |||||: : |    ||
Db      28 KYFFGKIGEGSQRSLTENKIRISKLKDIFLTGELNWSDIGGLPGMILTIADQGKSNLVLH 87

Qy     140 GPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPE---YEDETMTVYQIPI---HSEQRRGK 193
             :  |: :   :     ||:|    |   :   |:|: : |    :     | | |
Db      88 YGNDILNYIVSTWRYFVFRFGIDL--NDHIMKDKEVYKDKIIAVKSFNVLKNGGEDRLGV 145

Qy     194 HQPWQS------------PERPLSRLSPERSSDSESNENEPHLPHGVSQRRGVRDSSLVV 241
           :|                |  |  |  |||   |   | |                |: : |
Db     146 FDSFQKGVLRSIVAKMFPKHAPTDRYDP--SSDPHLNVELPDL-----------DAKVEV 192

Qy     242 AFICKLHLK--RGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSIT-HEGREILAEELCT 298
           :   ::     || | | :| ::|:| |     |: | :  |::||   |  :  |::
Db     193 STNYEISFSPVRGKFKVEEAIKLGVPKG-----PLFAKLTKGQTITLDNGIVVTPEQVLE 247

Qy     299 PPDPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPASVLVDSRYQQWM 358
             |  :::: ||: ::     ||  ::::    | : :| :     | ::     ::
Db     248 NERHFAKVLILDIPDDLYL-----NAFVEKFKDYDCAELGMVYYFLGDEVTINDNLFAFI 302

Qy     359 ERFGPDTQHLVLNENCASVHNLRSHKIQTQLNLIHPDIFPL----LTSFRCK-------- 406
           : |            |   |::: ||      | | |:         ||: : |
Db     303 DIFE--------KNNYGKVNHMISH------NKISPNTISFFGSALTTLKLKALQVNNYN 348

Qy     407 --KEGPTLS--------VPMVQGECLLKYQLRPRRE---------WQRDAIITCNP---- 443
             |    |          |: :|  : | |  |             : ::  :|  |
Db     349 LPKTDRVFSKDFYDRFDTPLSRGTSMCKSQEEPLNTIIEKDNIHIFSQNKTVTFEPFRMN 408

Qy     444 ---------------------EEFIVEALQLP--------NFQQSVQEYRRSAQDGPAPA 474
                               ||  |: |: |        | |  |  :  ||:
Db     409 EEPMKCNINGEVADFSWQEIFEEH-VKPLEFPLADVDTVINNQLHVDNFNNSAE------ 461

Qy     475 EKRSQYPEIIFLGTGSAIPMKIRNVSATLVNI--------SPDTSLLLDCGEGTFGQLCR 526
             : :: ||| |||||:| | ||| :||| :         : : :::|| || | | : |
Db     462 --KKKHVEIITLGTGSALPSKYRNVVSTLVKVPFTDADGNTINRNIMLDAGENTLGTIHR 519

Qy     527 HYGD-QVDRVLGTLAAVFVSHLHADHHTGLPSILLQRERALASLGKPLHPLLVVAPNQLK 585
           :    |  :   |  :::|||||||| |: |:|   |    :         : || |
Db     520 MFSQLAVKSIFQDLKMIYLSHLHADHHLGIISVL--NEWYKYNKDDETSYIYVVTP---- 573

Qy     586 AWLQQYHN--------QCQEVLHHISMIPA---------------------KCLQEGA- 614
           |  |||          : :|:|   |  |                       : |:| :
Db     574 -W--QYHKFVNEWLVLENKEILKRIKYISCEHFINDSFVRMQTQSVPLAEFNEILKENSN 630

Qy     615 -------EISSPAVER---LISSLLRTCDLEEFQTCLVRHCKHAFGCALV--------HT 656
                 |:   :  |   ||  :     :| ||||   ||  |:  ::          |
Db     631 QESNRKLELDRDSSYRDVDLIRQMYEDLSIEYFQTCRAIHCDWAYSNSITFRMDENNEHN 690

Qy     657 SGWKVVYSGDTMPC--EALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMR 714
           : :|| ||||| |   :  : :| :: |||||||||: | |:||:| | | ::|| |  :
Db     691 T-FKVSYSGDTRPNIEKFSLEIGYNSDLLIHEATLENQLLEDAVKKKHCTINEAIGVSNK 749

Qy     715 MNAEFIMLNHFSQRYAKVPLFSPN---FSEKVGVAFDHMKVCFGDFPTMPKLIPPLKALF 771
           |||  ::| |||||| |:|    |    :  :   ||| | | :       :: | |   |
Db     750 MNARKLILTHFSQRYPKLPQLDNNIDVMAREFCFAFDSMIVDYEKIGEQQRIFPLLNKAF 809

Qy     772 AGDIEEMEERREKRELRQVR 791
             :|| ||  :  ::  |:
Db     810 ---VEEKEEEEDVDDVESVQ 826
```

```
RESULT 9
US-09-564-805-229
; Sequence 229, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
;  APPLICANT:  Tavtigian, Sean V.
;  APPLICANT:  Teng, David H.F.
;  APPLICANT:  Simard, Jacques
;  APPLICANT:  Rommens, Johanna M.
;  APPLICANT:  Myriad Genetics, Inc.
;  TITLE OF INVENTION:  Chromosome 17p-Linked Prostate Cancer Susceptibility
;  TITLE OF INVENTION:  Gene and a Paralog and Orthologous Genes
;  FILE REFERENCE:  2318-258
;  CURRENT APPLICATION NUMBER: US/09/564,805
;  CURRENT FILING DATE:  2000-05-05
;  PRIOR APPLICATION NUMBER:  US 60/107,468
;  PRIOR FILING DATE:  1998-11-06
;  PRIOR APPLICATION NUMBER:  09/434,382
;  PRIOR FILING DATE:  1999-11-05
;  NUMBER OF SEQ ID NOS:  240
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 229
;   LENGTH: 838
;   TYPE: PRT
;   ORGANISM: Saccharomyces cerevisiae
US-09-564-805-229

  Query Match             13.9%;  Score 599.5;  DB 4;  Length 838;
  Best Local Similarity   25.7%;  Pred. No. 1.9e-49;
  Matches  221;  Conservative  138;  Mismatches  290;  Indels  211;  Gaps   36;

Qy      82 RYLF-NCGEGVQRLMQEHKLKVARLDNIFLT-RMHWSNVGGLSGMILTLKETGLPKCVLS 139
           :| |   ||| || : |:|:::::| :|||| ::||::||| |||||: : |    ||
Db      28 KYFFGKIGEGSQRSLTENKIRISKLKDIFLTGELNWSDIGGLPGMILTIADQGKSNLVLH 87

Qy     140 GPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPE---YEDETMTVYQIPI---HSEQRRGK 193
             :  |: :   :     ||:|    |   :   |:|: : |    :     | | |
Db      88 YGNDILNYIVSTWRYFVFRFGIDL--NDHIMKDKEVYKDKIIAVKSFNVLKNGGEDRLGV 145

Qy     194 HQPWQS------------PERPLSRLSPERSSDSESNENEPHLPHGVSQRRGVRDSSLVV 241
           :|                |  |  |  |||   |   | |                |: : |
Db     146 FDSFQKGVLRSIVAKMFPKHAPTDRYDP--SSDPHLNVELPDL-----------DAKVEV 192

Qy     242 AFICKLHLK--RGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSIT-HEGREILAEELCT 298
           :   ::     || | | :| ::|:| |     |: | :  |::||   |  :  |::
Db     193 STNYEISFSPVRGKFKVEEAIKLGVPKG-----PLFAKLTKGQTITLDNGIVVTPEQVLE 247

Qy     299 PPDPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPASVLVDSRYQQWM 358
             |  :::: ||: ::     ||  ::::    | : :| :     | ::     ::
Db     248 NERHFAKVLILDIPDDLYL-----NAFVEKFKDYDCAELGMVYYFLGDEVTINDNLFAFI 302

Qy     359 ERFGPDTQHLVLNENCASVHNLRSHKIQTQLNLIHPDIFPL----LTSFRCK-------- 406
           : |            |   |::: ||      | | |:         ||: : |
Db     303 DIFE--------KNNYGKVNHMISH------NKISPNTISFFGSALTTLKLKALQVNNYN 348

Qy     407 --KEGPTLS--------VPMVQGECLLKYQLRPRRE---------WQRDAIITCNP---- 443
             |    |          |: :|  : | |  |             : ::  :|  |
Db     349 LPKTDRVFSKDFYDRFDTPLSRGTSMCKSQEEPLNTIIEKDNIHIFSQNKTVTFEPFRMN 408

Qy     444 ---------------------EEFIVEALQLP--------NFQQSVQEYRRSAQDGPAPA 474
                               ||  |: |: |        | |  |  :  ||:
Db     409 EEPMKCNINGEVADFSWQEIFEEH-VKPLEFPLADVDTVINNQLHVDNFNNSAE------ 461

Qy     475 EKRSQYPEIIFLGTGSAIPMKIRNVSATLVNI--------SPDTSLLLDCGEGTFGQLCR 526
             : :: ||| |||||:| | ||| :||| :         : : :::|| || | | : |
Db     462 --KKKHVEIITLGTGSALPSKYRNVVSTLVKVPFTDADGNTINRNIMLDAGENTLGTIHR 519

Qy     527 HYGD-QVDRVLGTLAAVFVSHLHADHHTGLPSILLQRERALASLGKPLHPLLVVAPNQLK 585
           :    |  :   |  :::|||||||| |: |:|   |    :         : || |
Db     520 MFSQLAVKSIFQDLKMIYLSHLHADHHLGIISVL--NEWYKYNKDDETSYIYVVTP---- 573

Qy     586 AWLQQYHN--------QCQEVLHHISMIPA---------------------KCLQEGA- 614
```

```
Db    574 -W--QYHKFVNEWLVLENKEILKRIKYISCEHFINDSFVRMQTQSVPLAEFNEILKENSN 630

Qy    615 -------EISSPAVER---LISSLLRTCDLEEFQTCLVRHCKHAFGCALV--------HT 656
Db    631 QESNRKLELDRDSSYRDVDLIRQMYEDLSIEYFQTCRAIHCDWAYSNSITFRMDENNEHN 690

Qy    657 SGWKVVYSGDTMPC--EALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMR 714
Db    691 T-FKVSYSGDTRPNIEKFSLEIGYNSDLLIHEATLENQLLEDAVKKKHCTINEAIGVSNK 749

Qy    715 MNAEFIMLNHFSQRYAKVPLFSPN---FSEKVGVAFDHMKVCFGDFPTMPKLIPPLKALF 771
Db    750 MNARKLILTHFSQRYPKLPQLDNNIDVMAREFCFAFDSMIVDYEKIGEQQRIFPLLNKAF 809

Qy    772 AGDIEEMEERREKRELRQVR 791
Db    810 ---VEEKEEEEDVDDVESVQ 826
```

RESULT 10
US-09-564-805-211
; Sequence 211, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 211
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-564-805-211

Query Match 9.7%; Score 420; DB 4; Length 81;
Best Local Similarity 100.0%; Pred. No. 1.4e-33;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MWALCSLLRSAAGRTMSQGRTISQAPARRERPRKDPLRHLRTREKRGPSGCSGGPNTVYL 60
Db      1 MWALCSLLRSAAGRTMSQGRTISQAPARRERPRKDPLRHLRTREKRGPSGCSGGPNTVYL 60

Qy     61 QVVAAGSRDSGAALYVFSEFN 81
Db     61 QVVAAGSRDSGAALYVFSEFN 81
```

RESULT 11
US-09-564-805-232
; Sequence 232, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 232
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Methanobacterium thermoautotrophicum
US-09-564-805-232

Query Match 6.5%; Score 281; DB 4; Length 307;
Best Local Similarity 28.2%; Pred. No. 5.4e-19;
Matches 84; Conservative 45; Mismatches 93; Indels 76; Gaps 11;

```
Qy    482 EIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDCGEGTFGQLCRHYGDQVDRVLG---- 537
Db      3 EVTFLGTSSAVPSKNRNHTSIALRI-PGEIFLFDCGEGTQRQMA---------LAGISPM 52

Qy    538 TLAAVFVSHLHADHHTGLPSILLQ-----RERALASLGKP-LHPLLVVAPNQLKAWLQQY 591
Db     53 KVTRIFITHLHGDHILGIPGMIQSMGFRGREEPLDIYGPPGIHEL--------------- 97

Qy    592 HNQCQEVLHHISM---IPAKCLQEGAEI--------SSPAVERLISSLLRTCDLEEFQTC 640
Db     98 -HECIMKMGYFTLDFDINVHEVRGGTVVEEDDYRVTSAPASHSVFN--LAYCFEEKKRPR 154

Qy    641 LVRHCKHAFGC------ALVH--------------------TSGWKVVYSGDTMPCEAL 673
Db    155 FLREKAIALGLKPGPAFGKLHRGIPVRVGDRIIMPEEVLGSPRKGVKVCYSGDTRPCESV 214

Qy    674 VRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMRMNAEFIMLNHFSQRYAK 731
Db    215 IKLAEGAELLIHESTLEAGSEDKAAESGHSTAREAAEVARSAGVKRLILTHLSTRYKR 272
```

RESULT 12
US-09-564-805-213
; Sequence 213, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 213
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-564-805-213

Query Match 6.4%; Score 278; DB 4; Length 73;
Best Local Similarity 74.1%; Pred. No. 9.4e-20;
Matches 60; Conservative 2; Mismatches 11; Indels 8; Gaps 2;

```
Qy      1 MWALCSLLRSAAGRTMSQGRTISQAPARRERPRKDPLRHLRTREKRGPSGCSGGPNTVYL 60
Db      1 MWALRSLLRPLGLRTMSQG------SARRPRPPKDPLRHLRTREKRGPG--PGGPNTVYL 52
```

```
Qy      61 QVVAAGSRDSGAALYVFSEFN 81
           |||||| ||:||||||||||:|
Db      53 QVVAAGGRDAGAALYVFSEYN 73


RESULT 13
US-09-564-805-230
; Sequence 230, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
;  APPLICANT:  Tavtigian, Sean V.
;  APPLICANT:  Teng, David H.F.
;  APPLICANT:  Simard, Jacques
;  APPLICANT:  Rommens, Johanna M.
;  APPLICANT:  Myriad Genetics, Inc.
;  TITLE OF INVENTION:  Chromosome 17p-Linked Prostate Cancer Susceptibility
;  TITLE OF INVENTION:  Gene and a Paralog and Orthologous Genes
;  FILE REFERENCE:  2318-258
;  CURRENT APPLICATION NUMBER: US/09/564,805
;  CURRENT FILING DATE:  2000-05-05
;  PRIOR APPLICATION NUMBER:  US 60/107,468
;  PRIOR FILING DATE:  1998-11-06
;  PRIOR APPLICATION NUMBER:  09/434,382
;  PRIOR FILING DATE:  1999-11-05
;  NUMBER OF SEQ ID NOS:  240
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 230
;   LENGTH: 311
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-09-564-805-230

  Query Match             6.4%;  Score 275;  DB 4;  Length 311;
  Best Local Similarity  26.2%;  Pred. No. 2.1e-18;
  Matches   78;  Conservative   49;  Mismatches  109;  Indels   62;  Gaps    8;

Qy     476 KRSQYPEIIFLGTGSAIPMKIRNVSATLVNISPDTS---LLLDCGEGTFGQLCRHYGDQV 532
           || :  |:|||||  : :| : |||:| |:|:   |     | ||||||  ||     :
Db       2 KRDELMELIFLGTSAGVPTRTRNVTAILLNLQHPTQSGLWLFDCGEGTQHQLLHTAFNP- 60

Qy     533 DRVLGTLAAVFVSHLHADHHTGLPSILLQRERALASLGKPLHPLLVVAPNQLKAWLQ--- 589
           | |  :|:|||| ||  ||| :|  |     |:   : || :  |  :: :::
Db      61 ----GKLDKIFISHLHGDHLFGLPGLLCSR-----SMSGIIQPLTIYGPQGIREFVETAL 111

Qy     590 -------QYHNQCQEV---------LHHISMIP----------------------AKCLQ 611
                  |  :  |:         |  ::  |                       |: |:
Db     112 RISGSWTDYPLEIVEIGAGEILDDGLRKVTAYPLEHPLECYGYRIEEHDKPGALNAQALK 171

Qy     612 EGAEISSPAVERLISSLLRTCDLEEFQTCLVRHCKHAFGCALVHTSGWKVVYSGDTMPCE 671
           |  : | :   :|  ||: :        :          |  :   ||| ||:
Db     172 AAGVPPGPLFQELKAG--KTITLEDGRQI------NGADYLAAPVPGKALAIFGDTGPCD 223

Qy     672 ALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMRMNAEFIMLNHFSQRY 729
           | : : |   :::|||||:  :| :|  : ||:| || ::       ::: | | ||
Db     224 AALDLAKGVDVMVHEATLDITMEAKANSRGHSSTRQAATLAREAGVGKLIITHVSSRY 281


RESULT 14
US-09-564-805-220
; Sequence 220, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
;  APPLICANT:  Tavtigian, Sean V.
;  APPLICANT:  Teng, David H.F.
;  APPLICANT:  Simard, Jacques
;  APPLICANT:  Rommens, Johanna M.
;  APPLICANT:  Myriad Genetics, Inc.
;  TITLE OF INVENTION:  Chromosome 17p-Linked Prostate Cancer Susceptibility
;  TITLE OF INVENTION:  Gene and a Paralog and Orthologous Genes
;  FILE REFERENCE:  2318-258
;  CURRENT APPLICATION NUMBER: US/09/564,805
;  CURRENT FILING DATE:  2000-05-05
;  PRIOR APPLICATION NUMBER:  US 60/107,468
;  PRIOR FILING DATE:  1998-11-06
;  PRIOR APPLICATION NUMBER:  09/434,382
;  PRIOR FILING DATE:  1999-11-05
;  NUMBER OF SEQ ID NOS:  240
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 220
;   LENGTH: 363
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-564-805-220

  Query Match             5.7%;  Score 245.5;  DB 4;  Length 363;
  Best Local Similarity  26.4%;  Pred. No. 2.1e-15;
  Matches   87;  Conservative   42;  Mismatches  114;  Indels   87;  Gaps   10;

Qy     482 EIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDCGEGTFGQLCRHYGDQVDRVLGTLAA 541
           :: |||||:| |   |  || ::    :   | ||||||  || :       : | :
Db       4 DVTFLGTGAAYPSPTRGASAVVLRCEGE-CWLFDCGEGTQTQLMKS-----QLKAGRITK 57

Qy     542 VFVSHLHADHHTGLPSILLQRERALASLGKPLHPLLVVAPNQLKAWLQQ----------Y 591
           :|::||| ||  ||| :|        |:     |: :  |  |: :: :          :
Db      58 IFITHLHGDHFFGLPGLLCTISLQSGSMVSK-QPIEIYGPVGLRDFIWRTMELSHTELVF 116

Qy     592 HNQCQEVLHHISMIPAKCLQEGAEI----SSPAVERLISSLLRT-------CDLEEFQTC 640
           |   |::      ||: |:| | :    | |  |:  : || :        | |:|
Db     117 HYVVHELVPTADQCPAEELKEFAHVNRADSPPKEEQGRTILLDSEENSYLLFDDEQFVVK 176

Qy     641 LVR--HCKHAFGCALVHTS---------------------------------------- 657
             |  |   :|| ::|
Db     177 AFRLFHRIPSFGFSVVEKKRPGKLNAQKLKDLGVPPGPAYGKLKNGISVVLENGVTISPQ 236

Qy     658 --------GWKVVYSGDTMPCEALVRMGK-----DATLLIHEATLEDGLEEEAVEKTHST 704
                   | |:   ||   |  :|  |       :| ||||||||:|   ::| |  |||
Db     237 DVLKKPIVGRKICILGD---CSGVVGDGGVKLCFEADLLIHEATLDDAQMDKAKEHGHST 293

Qy     705 TSQAISVGMRMNAEFIMLNHFSQRYAKVPL 734
             | :      |: ::| ||||||  | |
Db     294 PQMAATFAKLCRAKRLVLTHFSQRYKPVAL 323


RESULT 15
US-09-564-805-231
; Sequence 231, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
;  APPLICANT:  Tavtigian, Sean V.
;  APPLICANT:  Teng, David H.F.
;  APPLICANT:  Simard, Jacques
;  APPLICANT:  Rommens, Johanna M.
;  APPLICANT:  Myriad Genetics, Inc.
;  TITLE OF INVENTION:  Chromosome 17p-Linked Prostate Cancer Susceptibility
;  TITLE OF INVENTION:  Gene and a Paralog and Orthologous Genes
;  FILE REFERENCE:  2318-258
;  CURRENT APPLICATION NUMBER: US/09/564,805
;  CURRENT FILING DATE:  2000-05-05
;  PRIOR APPLICATION NUMBER:  US 60/107,468
;  PRIOR FILING DATE:  1998-11-06
;  PRIOR APPLICATION NUMBER:  09/434,382
;  PRIOR FILING DATE:  1999-11-05
;  NUMBER OF SEQ ID NOS:  240
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 231
;   LENGTH: 326
;   TYPE: PRT
;   ORGANISM: Synechocystis sp.
US-09-564-805-231

  Query Match             5.6%;  Score 243.5;  DB 4;  Length 326;
  Best Local Similarity  26.2%;  Pred. No. 2.8e-15;
  Matches   84;  Conservative   46;  Mismatches  132;  Indels   59;  Gaps   10;

Qy     482 EIIFLGTGSAIPMKIRNVSATLVNISPDTSL-LLDCGEGTFGQLCRHYGDQVDRVLGTLA 540
```

```
          || |||| | :| : ||||:  : :     | | ||||||  |  |      :  :  |
Db      2 EITFLGTSSGVPTRNRNVSSIALRLPQRAELWLFDCGEGTQHQFLRS-----EVKISQLT 56

Qy    541 AVFVSHLHADHHTGLPSILLQRERALASLGKPLHPLLVVAPNQLKAWLQ----------- 589
          :|::||| ||  ||  :|      ||  |: :    :  |  |  :|:
Db     57 RIFITHLHGDHIFGLMGLL--ASSGLAGSGQGIE---IYGPEGLGDYLEACCRFSSTHLG 111

Qy    590 -----------------QYHNQCQEVLHHI---------SMIPAKCLQEGAEISSPAVE 622
                            :   |  : | |            | :   | ||
Db    112 KRLKVHTVRENGLIYEDKDFQVHCGLLKHRIPAYGYRVEEKQRPGRFNVEQAEALGIPFG 171

Qy    623 RLISSLLRTCDLEEFQTCLVRHCKHAFGCALVH--TSGWKVVYSGDTMPCEALVRMGKDA 680
          :         |:: :|  :   :   |  |      | | ||  ||: ||  : : ::|
Db    172 PIYG------QLKQGKTVTLEDGRRIRGQDLCEPPEPGRKFVYCTDTVFCEEAIALAQEA 225

Qy    681 TLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMRMNAEFIMLNHFSQRYAK-VPLFSPNF 739
          ||:||||      : | :: |||:: |  | :  | : ::: ||| |||   ||   |
Db    226 DLLVHEATFAHQDAQLAFDRLHSTSTMAAQVALLANVKQLIMTHFSPRYAPGNPLQLENL 285

Qy    740 SEKVGVAFDHMKVCFGDFPTM 760
            :   | : ::   || |:
Db    286 LAEAQAIFPNTRLA-RDFLTV 305
```

Search completed: May 14, 2003, 10:07:00
Job time : 21 secs

OM protein - protein search, using sw model

Run on: May 14, 2003, 10:05:56 ; Search time 26 Seconds
(without alignments)
2923.586 Million cell updates/sec

Title: US-09-434-382-2
Perfect score: 4325
Sequence: 1 MWALCSLLRSAAGRTMSQGR.....EPQQKRAHTEEPQAKKVRAQ 826

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 349150 seqs, 92025710 residues

Total number of hits satisfying chosen parameters: 349150

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4325	100.0	826	9	US-09-988-626-2	Sequence 2, Appli
2	4325	100.0	826	9	US-09-988-687-2	Sequence 2, Appli
3	4283	99.0	826	9	US-09-988-626-224	Sequence 224, App
4	4283	99.0	826	9	US-09-988-687-224	Sequence 224, App
5	4261	98.5	826	9	US-09-988-626-226	Sequence 226, App
6	4261	98.5	826	9	US-09-988-687-226	Sequence 226, App
7	3473.5	80.3	822	9	US-09-988-626-222	Sequence 222, App
8	3473.5	80.3	822	9	US-09-988-687-222	Sequence 222, App
9	875.5	20.2	837	9	US-09-988-626-228	Sequence 228, App
10	875.5	20.2	837	9	US-09-988-687-228	Sequence 228, App
11	760	17.6	844	9	US-09-988-626-227	Sequence 227, App
12	760	17.6	844	9	US-09-988-687-227	Sequence 227, App
13	599.5	13.9	838	9	US-09-988-626-229	Sequence 229, App
14	599.5	13.9	838	9	US-09-988-687-229	Sequence 229, App
15	420	9.7	81	9	US-09-988-626-211	Sequence 211, App
16	420	9.7	81	9	US-09-988-687-211	Sequence 211, App
17	281	6.5	307	9	US-09-988-626-232	Sequence 232, App
18	281	6.5	307	9	US-09-988-687-232	Sequence 232, App
19	278	6.4	73	9	US-09-988-626-213	Sequence 213, App
20	278	6.4	73	9	US-09-988-687-213	Sequence 213, App
21	275	6.4	311	9	US-09-988-626-230	Sequence 230, App
22	275	6.4	311	9	US-09-988-687-230	Sequence 230, App
23	245.5	5.7	363	9	US-09-988-626-220	Sequence 220, App
24	245.5	5.7	363	9	US-09-988-687-220	Sequence 220, App
25	243.5	5.6	326	9	US-09-988-626-231	Sequence 231, App
26	243.5	5.6	326	9	US-09-988-687-231	Sequence 231, App
27	142.5	3.3	166	10	US-09-925-301-1076	Sequence 1076, Ap
28	115.5	2.7	255	9	US-09-738-626-6252	Sequence 6252, Ap
29	115	2.7	1400	9	US-10-123-036-4	Sequence 4, Appli
30	112	2.6	1404	10	US-09-862-027-24	Sequence 24, Appl
31	111.5	2.6	1054	10	US-09-798-042-87	Sequence 87, Appl
32	110.5	2.6	381	9	US-09-764-868-915	Sequence 915, App
33	104	2.4	782	9	US-09-908-193-47	Sequence 47, Appl
34	102.5	2.4	1243	9	US-10-196-935A-4	Sequence 4, Appli
35	102	2.4	1356	9	US-09-969-037-7	Sequence 7, Appli
36	102	2.4	1356	9	US-10-022-939-2	Sequence 2, Appli
37	102	2.4	1356	9	US-10-100-405A-2	Sequence 2, Appli
38	99.5	2.3	1142	9	US-10-085-108-7	Sequence 7, Appli
39	99.5	2.3	1142	10	US-09-899-651-2	Sequence 2, Appli
40	99	2.3	896	10	US-09-923-563A-1	Sequence 1, Appli
41	98.5	2.3	847	10	US-09-476-242-2	Sequence 2, Appli
42	98	2.3	896	9	US-10-206-566-3	Sequence 3, Appli
43	98	2.3	1029	9	US-10-033-245-22	Sequence 22, Appl
44	98	2.3	1029	9	US-10-033-223-22	Sequence 22, Appl
45	98	2.3	1029	9	US-10-033-167-22	Sequence 22, Appl

ALIGNMENTS

RESULT 1
US-09-988-626-2
; Sequence 2, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-988-626-2

Query Match 100.0%; Score 4325; DB 9; Length 826;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 826; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MWALCSLLRSAAGRTMSQGRTISQAPARRERPRKDPLRHLRTREKRGPSGCSGGPNTVYL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MWALCSLLRSAAGRTMSQGRTISQAPARRERPRKDPLRHLRTREKRGPSGCSGGPNTVYL 60

Qy      61 QVVAAGSRDSGAALYVFSEFNRYLFNCGEGVQRLMQEHKLKVARLDNIFLTRMHWSNVGG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 QVVAAGSRDSGAALYVFSEFNRYLFNCGEGVQRLMQEHKLKVARLDNIFLTRMHWSNVGG 120

Qy     121 LSGMILTLKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPEYEDETMTV 180
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 LSGMILTLKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPEYEDETMTV 180

Qy     181 YQIPIHSEQRRGKHQPWQSPERPLSRLSPERSSDSESNENEPHLPHGVSQRRGVRDSSLV 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 YQIPIHSEQRRGKHQPWQSPERPLSRLSPERSSDSESNENEPHLPHGVSQRRGVRDSSLV 240

Qy     241 VAFICKLHLKRGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSITHEGREILAEELCTPP 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 VAFICKLHLKRGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSITHEGREILAEELCTPP 300

Qy     301 DPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPASVLVDSRYQQWMER 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 DPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPASVLVDSRYQQWMER 360

Qy     361 FGPDTQHLVLNENCASVHNLRSHKIQTQLNLIHPDIFPLLTSFRCKKEGPTLSVPMVQGE 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 FGPDTQHLVLNENCASVHNLRSHKIQTQLNLIHPDIFPLLTSFRCKKEGPTLSVPMVQGE 420

Qy     421 CLLKYQLRPRREWQRDAIITCNPEEFIVEALQLPNFQQSVQEYRRSAQDGPAPAEKRSQY 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CLLKYQLRPRREWQRDAIITCNPEEFIVEALQLPNFQQSVQEYRRSAQDGPAPAEKRSQY 480

Qy     481 PEIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDCGEGTFGQLCRHYGDQVDRVLGTLA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 PEIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDCGEGTFGQLCRHYGDQVDRVLGTLA 540

Qy     541 AVFVSHLHADHHTGLPSILLQRERALASLGKPLHPLLVVAPNQLKAWLQQYHNQCQEVLH 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AVFVSHLHADHHTGLPSILLQRERALASLGKPLHPLLVVAPNQLKAWLQQYHNQCQEVLH 600

Qy     601 HISMIPAKCLQEGAEISSPAVERLISSLLRTCDLEEFQTCLVRHCKHAFGCALVHTSGWK 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 HISMIPAKCLQEGAEISSPAVERLISSLLRTCDLEEFQTCLVRHCKHAFGCALVHTSGWK 660

Qy     661 VVYSGDTMPCEALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMRMNAEFI 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 VVYSGDTMPCEALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMRMNAEFI 720

Qy     721 MLNHFSQRYAKVPLFSPNFSEKVGVAFDHMKVCFGDFPTMPKLIPPLKALFAGDIEEMEE 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 MLNHFSQRYAKVPLFSPNFSEKVGVAFDHMKVCFGDFPTMPKLIPPLKALFAGDIEEMEE 780

Qy     781 RREKRELRQVRAALLSRELAGGLEDGEPQQKRAHTEEPQAKKVRAQ 826
           |||||||||||||||||||||||||||||||||||||||||||||
Db     781 RREKRELRQVRAALLSRELAGGLEDGEPQQKRAHTEEPQAKKVRAQ 826


RESULT 2
US-09-988-687-2
; Sequence 2, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
;  APPLICANT: Tavtigian, Sean V.
;  APPLICANT:  Teng, David H.F.
;  APPLICANT:  Simard, Jacques
;  APPLICANT:  Rommens, Johanna M.
;  APPLICANT:  Myriad Genetics, Inc.
;  TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
;  TITLE OF INVENTION:  Gene and a Paralog and Orthologous Genes
;  FILE REFERENCE: 2318-258
;  CURRENT APPLICATION NUMBER: US/09/988,687
;  CURRENT FILING DATE:  2001-11-20
;  PRIOR APPLICATION NUMBER: 09/564,805
;  PRIOR FILING DATE: 2000-05-05
;  PRIOR APPLICATION NUMBER: US 60/107,468
;  PRIOR FILING DATE: 1998-11-06
;  PRIOR APPLICATION NUMBER: 09/434,382
;  PRIOR FILING DATE: 1999-11-05
;  NUMBER OF SEQ ID NOS: 240
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 826
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-988-687-2

  Query Match           100.0%;  Score 4325;  DB 9;  Length 826;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  826;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MWALCSLLRSAAGRTMSQGRTISQAPARRERPRKDPLRHLRTREKRGPSGCSGGPNTVYL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MWALCSLLRSAAGRTMSQGRTISQAPARRERPRKDPLRHLRTREKRGPSGCSGGPNTVYL 60

Qy      61 QVVAAGSRDSGAALYVFSEFNRYLFNCGEGVQRLMQEHKLKVARLDNIFLTRMHWSNVGG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 QVVAAGSRDSGAALYVFSEFNRYLFNCGEGVQRLMQEHKLKVARLDNIFLTRMHWSNVGG 120

Qy     121 LSGMILTLKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPEYEDETMTV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 LSGMILTLKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPEYEDETMTV 180

Qy     181 YQIPIHSEQRRGKHQPWQSPERPLSRLSPERSSDSESNENEPHLPHGVSQRRGVRDSSLV 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 YQIPIHSEQRRGKHQPWQSPERPLSRLSPERSSDSESNENEPHLPHGVSQRRGVRDSSLV 240

Qy     241 VAFICKLHLKRGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSITHEGREILAEELCTPP 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 VAFICKLHLKRGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSITHEGREILAEELCTPP 300

Qy     301 DPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPASVLVDSRYQQWMER 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 DPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPASVLVDSRYQQWMER 360

Qy     361 FGPDTQHLVLNENCASVHNLRSHKIQTQLNLIHPDIFPLLTSFRCKKEGPTLSVPMVQGE 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 FGPDTQHLVLNENCASVHNLRSHKIQTQLNLIHPDIFPLLTSFRCKKEGPTLSVPMVQGE 420

Qy     421 CLLKYQLRPRREWQRDAIITCNPEEFIVEALQLPNFQQSVQEYRRSAQDGPAPAEKRSQY 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CLLKYQLRPRREWQRDAIITCNPEEFIVEALQLPNFQQSVQEYRRSAQDGPAPAEKRSQY 480

Qy     481 PEIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDCGEGTFGQLCRHYGDQVDRVLGTLA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 PEIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDCGEGTFGQLCRHYGDQVDRVLGTLA 540

Qy     541 AVFVSHLHADHHTGLPSILLQRERALASLGKPLHPLLVVAPNQLKAWLQQYHNQCQEVLH 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AVFVSHLHADHHTGLPSILLQRERALASLGKPLHPLLVVAPNQLKAWLQQYHNQCQEVLH 600

Qy     601 HISMIPAKCLQEGAEISSPAVERLISSLLRTCDLEEFQTCLVRHCKHAFGCALVHTSGWK 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 HISMIPAKCLQEGAEISSPAVERLISSLLRTCDLEEFQTCLVRHCKHAFGCALVHTSGWK 660

Qy     661 VVYSGDTMPCEALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMRMNAEFI 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 VVYSGDTMPCEALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMRMNAEFI 720

Qy     721 MLNHFSQRYAKVPLFSPNFSEKVGVAFDHMKVCFGDFPTMPKLIPPLKALFAGDIEEMEE 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 MLNHFSQRYAKVPLFSPNFSEKVGVAFDHMKVCFGDFPTMPKLIPPLKALFAGDIEEMEE 780

Qy     781 RREKRELRQVRAALLSRELAGGLEDGEPQQKRAHTEEPQAKKVRAQ 826
           |||||||||||||||||||||||||||||||||||||||||||||
Db     781 RREKRELRQVRAALLSRELAGGLEDGEPQQKRAHTEEPQAKKVRAQ 826


RESULT 3
US-09-988-626-224
; Sequence 224, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
;  APPLICANT: Tavtigian, Sean V.
;  APPLICANT:  Teng, David H.F.
;  APPLICANT:  Simard, Jacques
```

```
;  APPLICANT:  Rommens, Johanna M.
;  APPLICANT:  Myriad Genetics, Inc.
;  TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
;  TITLE OF INVENTION:  Gene and a Paralog and Orthologous Genes
;  FILE REFERENCE: 2318-258
;  CURRENT APPLICATION NUMBER: US/09/988,626
;  CURRENT FILING DATE:  2001-11-20
;  PRIOR APPLICATION NUMBER: 09/564,805
;  PRIOR FILING DATE: 2000-05-05
;  PRIOR APPLICATION NUMBER: US 60/107,468
;  PRIOR FILING DATE: 1998-11-06
;  PRIOR APPLICATION NUMBER: 09/434,382
;  PRIOR FILING DATE: 1999-11-05
;  NUMBER OF SEQ ID NOS: 240
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 224
;   LENGTH: 826
;   TYPE: PRT
;   ORGANISM: Pan troglodytes
US-09-988-626-224

  Query Match             99.0%;  Score 4283;  DB 9;  Length 826;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches  817;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 MWALCSLLRSAAGRTMSQGRTISQAPARRERPRKDPLRHLRTREKRGPSGCSGGPNTVYL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MWALCSLLRSAAGRTMSQGRTISQAPARRERPRKDPLRHLRTREKRGPSGCSGGPNTVYL 60

Qy     61 QVVAAGSRDSGAALYVFSEFNRYLFNCGEGVQRLMQEHKLKVARLDNIFLTRMHWSNVGG 120
          ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db     61 QVVAAGSRDSGAALYVFSEFNRYLFNCGEGIQRLMQEHKLKVARLDNIFLTRMHWSNVGG 120

Qy    121 LSGMILTLKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPEYEDETMTV 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 LSGMILTLKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPEYEDETMTV 180

Qy    181 YQIPIHSEQRRGKHQPWQSPERPLSRLSPERSSDSESNENEPHLPHGVSQRRGVRDSSLV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 YQIPIHSEQRRGKHQPWQSPERPLSRLSPERSSDSESNENEPHLPHGVSQRRGVRDSSLV 240

Qy    241 VAFICKLHLKRGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSITHEGREILAEELCTPP 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 VAFICKLHLKRGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSITHEGREILAEELCTPP 300

Qy    301 DPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPASVLVDSRYQQWMER 360
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    301 DPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPESVLVDSRYQQWMER 360

Qy    361 FGPDTQHLVLNENCASVHNLRSHKIQTQLNLIHPDIFPLLTSFRCKKEGPTLSVPMVQGE 420
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    361 FGPDTQHLVLNENCASVHNLRSHKIQTQLNLIHPDIFPLLTSFPCKKEGPTLSVPMVQGE 420

Qy    421 CLLKYQLRPRREWQRDAIITCNPEEFIVEALQLPNFQQSVQEYRRSAQDGPAPAEKRSQY 480
          ||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db    421 CLLKYQLRPRREWQRDAIITCNPEEFIIEALQLPNFQQSVQEYRRSAQDGPAPAEKRSQY 480

Qy    481 PEIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDCGEGTFGQLCRHYGDQVDRVLGTLA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 PEIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDCGEGTFGQLCRHYGDQVDRVLGTLA 540

Qy    541 AVFVSHLHADHHTGLPSILLQRERALASLGKPLHPLLVVAPNQLKAWLQQYHNQCQEVLH 600
          |||||||||||||| :||||||||||||| ||||||||||||||||||||||||||||||
Db    541 AVFVSHLHADHHTGLLNILLQRERALASLGKPFHPLLVVAPNQLKAWLQQYHNQCQEVLH 600

Qy    601 HISMIPAKCLQEGAEISSPAVERLISSLLRTCDLEEFQTCLVRHCKHAFGCALVHTSGWK 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 HISMIPAKCLQEGAEISSPAVERLISSLLRTCDLEEFQTCLVRHCKHAFGCALVHTSGWK 660

Qy    661 VVYSGDTMPCEALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMRMNAEFI 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 VVYSGDTMPCEALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMRMNAEFI 720

Qy    721 MLNHFSQRYAKVPLFSPNFSEKVGVAFDHMKVCFGDFPTMPKLIPPLKALFAGDIEEMEE 780
          ||||||||||||||||||:||||||||||||||| ||||||||||||||||||||||||||
Db    721 MLNHFSQRYAKVPLFSPNFNEKVGVAFDHMKVCFGDFATMPKLIPPLKALFAGDIEEMEE 780

Qy    781 RREKRELRQVRAALLSRELAGGLEDGEPQQKRAHTEEPQAKKVRAQ 826
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    781 RREKRELRQVRAALLSRELAGGLEDGEPQQKRAHTEEPQAKKVRAQ 826


RESULT 4
US-09-988-687-224
; Sequence 224, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
;  APPLICANT: Tavtigian, Sean V.
;  APPLICANT:  Teng, David H.F.
;  APPLICANT:  Simard, Jacques
;  APPLICANT:  Rommens, Johanna M.
;  APPLICANT:  Myriad Genetics, Inc.
;  TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
;  TITLE OF INVENTION:  Gene and a Paralog and Orthologous Genes
;  FILE REFERENCE: 2318-258
;  CURRENT APPLICATION NUMBER: US/09/988,687
;  CURRENT FILING DATE:  2001-11-20
;  PRIOR APPLICATION NUMBER: 09/564,805
;  PRIOR FILING DATE: 2000-05-05
;  PRIOR APPLICATION NUMBER: US 60/107,468
;  PRIOR FILING DATE: 1998-11-06
;  PRIOR APPLICATION NUMBER: 09/434,382
;  PRIOR FILING DATE: 1999-11-05
;  NUMBER OF SEQ ID NOS: 240
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 224
;   LENGTH: 826
;   TYPE: PRT
;   ORGANISM: Pan troglodytes
US-09-988-687-224

  Query Match             99.0%;  Score 4283;  DB 9;  Length 826;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches  817;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 MWALCSLLRSAAGRTMSQGRTISQAPARRERPRKDPLRHLRTREKRGPSGCSGGPNTVYL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MWALCSLLRSAAGRTMSQGRTISQAPARRERPRKDPLRHLRTREKRGPSGCSGGPNTVYL 60

Qy     61 QVVAAGSRDSGAALYVFSEFNRYLFNCGEGVQRLMQEHKLKVARLDNIFLTRMHWSNVGG 120
          ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db     61 QVVAAGSRDSGAALYVFSEFNRYLFNCGEGIQRLMQEHKLKVARLDNIFLTRMHWSNVGG 120

Qy    121 LSGMILTLKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPEYEDETMTV 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 LSGMILTLKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPEYEDETMTV 180

Qy    181 YQIPIHSEQRRGKHQPWQSPERPLSRLSPERSSDSESNENEPHLPHGVSQRRGVRDSSLV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 YQIPIHSEQRRGKHQPWQSPERPLSRLSPERSSDSESNENEPHLPHGVSQRRGVRDSSLV 240

Qy    241 VAFICKLHLKRGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSITHEGREILAEELCTPP 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 VAFICKLHLKRGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSITHEGREILAEELCTPP 300

Qy    301 DPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPASVLVDSRYQQWMER 360
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    301 DPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPESVLVDSRYQQWMER 360

Qy    361 FGPDTQHLVLNENCASVHNLRSHKIQTQLNLIHPDIFPLLTSFRCKKEGPTLSVPMVQGE 420
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    361 FGPDTQHLVLNENCASVHNLRSHKIQTQLNLIHPDIFPLLTSFPCKKEGPTLSVPMVQGE 420

Qy    421 CLLKYQLRPRREWQRDAIITCNPEEFIVEALQLPNFQQSVQEYRRSAQDGPAPAEKRSQY 480
          ||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db    421 CLLKYQLRPRREWQRDAIITCNPEEFIIEALQLPNFQQSVQEYRRSAQDGPAPAEKRSQY 480
```

```
Qy    481 PEIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDCGEGTFGQLCRHYGDQVDRVLGTLA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 PEIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDCGEGTFGQLCRHYGDQVDRVLGTLA 540

Qy    541 AVFVSHLHADHHTGLPSILLQRERALASLGKPLHPLLVVAPNQLKAWLQQYHNQCQEVLH 600
          |||||||||||||| :|||||||||||||| ||||||||||||||||||||||||||||||
Db    541 AVFVSHLHADHHTGLLNILLQRERALASLGKPFHPLLVVAPNQLKAWLQQYHNQCQEVLH 600

Qy    601 HISMIPAKCLQEGAEISSPAVERLISSLLRTCDLEEFQTCLVRHCKHAFGCALVHTSGWK 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 HISMIPAKCLQEGAEISSPAVERLISSLLRTCDLEEFQTCLVRHCKHAFGCALVHTSGWK 660

Qy    661 VVYSGDTMPCEALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMRMNAEFI 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 VVYSGDTMPCEALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMRMNAEFI 720

Qy    721 MLNHFSQRYAKVPLFSPNFSEKVGVAFDHMKVCFGDFPTMPKLIPPLKALFAGDIEEMEE 780
          |||||||||||||||||:|||||||||||||| ||||||||||||||||||||||||||||
Db    721 MLNHFSQRYAKVPLFSPNFNEKVGVAFDHMKVCFGDFATMPKLIPPLKALFAGDIEEMEE 780

Qy    781 RREKRELRQVRAALLSRELAGGLEDGEPQQKRAHTEEPQAKKVRAQ 826
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    781 RREKRELRQVRAALLSRELAGGLEDGEPQQKRAHTEEPQAKKVRAQ 826
```

RESULT 5
US-09-988-626-226
; Sequence 226, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 226
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Gorilla gorilla
US-09-988-626-226

Query Match 98.5%; Score 4261; DB 9; Length 826;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 814; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

```
Qy      1 MWALCSLLRSAAGRTMSQGRTISQAPARRERPRKDPLRHLRTREKRGPSGCSGGPNTVYL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MWALCSLLRSAAGRTMSQGRTISQAPARRERPRKDPLRHLRTREKRGPSGCSGGPNTVYL 60

Qy     61 QVVAAGSRDSGAALYVFSEFNRYLFNCGEGVQRLMQEHKLKVARLDNIFLTRMHWSNVGG 120
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db     61 QVVAAGSRDSGAALYVFSEFNRYLFNCGEGVQRLMQEHKLKVVRLDNIFLTRMHWSNVGG 120

Qy    121 LSGMILTLKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPEYEDETMTV 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 LSGMILTLKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPEYEDETMTV 180

Qy    181 YQIPIHSEQRRGKHQPWQSPERPLSRLSPERSSDSESNENEPHLPHGVSQRRGVRDSSLV 240
          ||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db    181 YQIPIHSEQRRGRHQPWQSPERPLSRLSPERSSDSESNENEPHLPHGVSQRRGVRDSSLV 240

Qy    241 VAFICKLHLKRGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSITHEGREILAEELCTPP 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 VAFICKLHLKRGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSITHEGREILAEELCTPP 300

Qy    301 DPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPASVLVDSRYQQWMER 360
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    301 DPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPESVLVDSRYQQWMER 360

Qy    361 FGPDTQHLVLNENCASVHNLRSHKIQTQLNLIHPDIFPLLTSFRCKKEGPTLSVPMVQGE 420
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    361 FGPDTQHLVLNENCASVHNLRSHKIQTQLNLIHPDIFPLLTSFPCKKEGPTLSVPMVQGE 420

Qy    421 CLLKYQLRPRREWQRDAIITCNPEEFIVEALQLPNFQQSVQEYRRSAQDGPAPAEKRSQY 480
          |||||||||||||||||||||||||||||||||||||||||||| || ||||||||||||||
Db    421 CLLKYQLRPRREWQRDAIITCNPEEFIVEALQLPNFQQSVQEYRRSVQDVPAPAEKRSQY 480

Qy    481 PEIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDCGEGTFGQLCRHYGDQVDRVLGTLA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 PEIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDCGEGTFGQLCRHYGDQVDRVLGTLA 540

Qy    541 AVFVSHLHADHHTGLPSILLQRERALASLGKPLHPLLVVAPNQLKAWLQQYHNQCQEVLH 600
          |||||||||||||| :|||||:|||||||||||||||||:|||||||||||||||||||||
Db    541 AVFVSHLHADHHTGLLNILLQREQALASLGKPLHPLLVVAPSQLKAWLQQYHNQCQEVLH 600

Qy    601 HISMIPAKCLQEGAEISSPAVERLISSLLRTCDLEEFQTCLVRHCKHAFGCALVHTSGWK 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 HISMIPAKCLQEGAEISSPAVERLISSLLRTCDLEEFQTCLVRHCKHAFGCALVHTSGWK 660

Qy    661 VVYSGDTMPCEALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMRMNAEFI 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 VVYSGDTMPCEALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMRMNAEFI 720

Qy    721 MLNHFSQRYAKVPLFSPNFSEKVGVAFDHMKVCFGDFPTMPKLIPPLKALFAGDIEEMEE 780
          |||||||||||||||||:|||||||||||||| ||||||||||||||||||||||||||||
Db    721 MLNHFSQRYAKVPLFSPNFNEKVGVAFDHMKVCFGDFPTMPKLIPPLKALFAGDIEEMEE 780

Qy    781 RREKRELRQVRAALLSRELAGGLEDGEPQQKRAHTEEPQAKKVRAQ 826
          |||||||||||||||| ||||||||||||||||||||||||||||||
Db    781 RREKRELRQVRAALLSGELAGGLEDGEPQQKRAHTEEPQAKKVRAQ 826
```

RESULT 6
US-09-988-687-226
; Sequence 226, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 226
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Gorilla gorilla
US-09-988-687-226

Query Match 98.5%; Score 4261; DB 9; Length 826;

```
  Best Local Similarity   98.5%;  Pred. No. 0;
  Matches  814;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 MWALCSLLRSAAGRTMSQGRTISQAPARRERPRKDPLRHLRTREKRGPSGCSGGPNTVYL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MWALCSLLRSAAGRTMSQGRTISQAPARRERPRKDPLRHLRTREKRGPSGCSGGPNTVYL 60

Qy      61 QVVAAGSRDSGAALYVFSEFNRYLFNCGEGVQRLMQEHKLKVARLDNIFLTRMHWSNVGG 120
           |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db      61 QVVAAGSRDSGAALYVFSEFNRYLFNCGEGVQRLMQEHKLKVVRLDNIFLTRMHWSNVGG 120

Qy     121 LSGMILTLKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPEYEDETMTV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 LSGMILTLKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPEYEDETMTV 180

Qy     181 YQIPIHSEQRRGKHQPWQSPERPLSRLSPERSSDSESNENEPHLPHGVSQRRGVRDSSLV 240
           ||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 YQIPIHSEQRRGRHQPWQSPERPLSRLSPERSSDSESNENEPHLPHGVSQRRGVRDSSLV 240

Qy     241 VAFICKLHLKRGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSITHEGREILAEELCTPP 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 VAFICKLHLKRGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSITHEGREILAEELCTPP 300

Qy     301 DPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPASVLVDSRYQQWMER 360
           |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     301 DPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPESVLVDSRYQQWMER 360

Qy     361 FGPDTQHLVLNENCASVHNLRSHKIQTQLNLIHPDIFPLLTSFRCKKEGPTLSVPMVQGE 420
           |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     361 FGPDTQHLVLNENCASVHNLRSHKIQTQLNLIHPDIFPLLTSFPCKKEGPTLSVPMVQGE 420

Qy     421 CLLKYQLRPRREWQRDAIITCNPEEFIVEALQLPNFQQSVQEYRRSAQDGPAPAEKRSQY 480
           |||||||||||||||||||||||||||||||||||||||||||| || ||||||||||||||
Db     421 CLLKYQLRPRREWQRDAIITCNPEEFIVEALQLPNFQQSVQEYRRSVQDVPAPAEKRSQY 480

Qy     481 PEIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDCGEGTFGQLCRHYGDQVDRVLGTLA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 PEIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDCGEGTFGQLCRHYGDQVDRVLGTLA 540

Qy     541 AVFVSHLHADHHTGLPSILLQRERALASLGKPLHPLLVVAPNQLKAWLQQYHNQCQEVLH 600
           ||||||||||||||| :||||||:||||||||||||||:||||||||||||||||||||||
Db     541 AVFVSHLHADHHTGLLNILLQREQALASLGKPLHPLLVVAPSQLKAWLQQYHNQCQEVLH 600

Qy     601 HISMIPAKCLQEGAEISSPAVERLISSLLRTCDLEEFQTCLVRHCKHAFGCALVHTSGWK 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 HISMIPAKCLQEGAEISSPAVERLISSLLRTCDLEEFQTCLVRHCKHAFGCALVHTSGWK 660

Qy     661 VVYSGDTMPCEALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMRMNAEFI 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 VVYSGDTMPCEALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMRMNAEFI 720

Qy     721 MLNHFSQRYAKVPLFSPNFSEKVGVAFDHMKVCFGDFPTMPKLIPPLKALFAGDIEEMEE 780
           ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db     721 MLNHFSQRYAKVPLFSPNFNEKVGVAFDHMKVCFGDFPTMPKLIPPLKALFAGDIEEMEE 780

Qy     781 RREKRELRQVRAALLSRELAGGLEDGEPQQKRAHTEEPQAKKVRAQ 826
           |||||||||||||||| ||||||||||||||||||||||||||||||
Db     781 RREKRELRQVRAALLSGELAGGLEDGEPQQKRAHTEEPQAKKVRAQ 826


RESULT 7
US-09-988-626-222
; Sequence 222, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
;  APPLICANT: Tavtigian, Sean V.
;  APPLICANT:  Teng, David H.F.
;  APPLICANT:  Simard, Jacques
;  APPLICANT:  Rommens, Johanna M.
;  APPLICANT:  Myriad Genetics, Inc.
;  TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
;  TITLE OF INVENTION:  Gene and a Paralog and Orthologous Genes
;  FILE REFERENCE: 2318-258
;  CURRENT APPLICATION NUMBER: US/09/988,626
;  CURRENT FILING DATE:  2001-11-20
;  PRIOR APPLICATION NUMBER: 09/564,805
;  PRIOR FILING DATE: 2000-05-05
;  PRIOR APPLICATION NUMBER: US 60/107,468
;  PRIOR FILING DATE: 1998-11-06
;  PRIOR APPLICATION NUMBER: 09/434,382
;  PRIOR FILING DATE: 1999-11-05
;  NUMBER OF SEQ ID NOS: 240
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 222
;   LENGTH: 822
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-988-626-222

  Query Match             80.3%;  Score 3473.5;  DB 9;  Length 822;
  Best Local Similarity   80.5%;  Pred. No. 6.9e-294;
  Matches  665;  Conservative   66;  Mismatches   76;  Indels   19;  Gaps    6;

Qy       1 MWALCSLLRSAAGRTMSQGRTISQAPARRERPRKDPLRHLRTREKRGPSGCSGGPNTVYL 60
           |||| ||||    ||||||        ||| || ||||||||||||||    ||||||||
Db       1 MWALRSLLRPLGLRTMSQG------SARRPRPPKDPLRHLRTREKRGPG--PGGPNTVYL 52

Qy      61 QVVAAGSRDSGAALYVFSEFNRYLFNCGEGVQRLMQEHKLKVARLDNIFLTRMHWSNVGG 120
           |||||| ||:|||||||||:|||||||||||||||||| :   ||||||||||||||||
Db      53 QVVAAGGRDAGAALYVFSEYNRYLFNCGEGVQRLMQEHKTESRSLDNIFLTRMHWSNVGG 112

Qy     121 LSGMILTLKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPEYEDETMTV 180
           | ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db     113 LCGMILTLKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPEYKDETMTV 172

Qy     181 YQIPIHSEQRRGKHQPWQSPERPLSRLSPERSSDSESNENEPHLPHGVSQRRGV-RDSSL 239
           ||:|||||:| || || |||    :||||::|||| | ||      |  |:  : :  ||
Db     173 YQVPIHSERRCGKQQPSQSPRTSPNRLSPKQSSDSGSAEN------GQCQQESMGQGPSL 226

Qy     240 VVAFICKLHLKRGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSITYEGREIAAEELCTP 299
           ||||:|||||::|||||||||:|||||||||||||||||||||:||||| |||||||
Db     227 VVAFVCKLHLRKGNFLVLKAKELGLPVGTAAIAPIIAAVKDGKSITYEGREIAAEELCTP 286

Qy     300 PDPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPASVLVDSRYQQWME 359
           ||||  |:|||||| || ||||| ||:||| :||||||||||:|| |||:|||||||||
Db     287 PDPGLVFIVVECPDEGFILPICENDTFKRYQAEADAPVALVVHIAPESVLIDSRYQQWME 346

Qy     360 RFGPDTQHLVLNENCASVHNLRSHKIQTQLNLIHPDIFPLLTSFRCKKEGPTLSVPMVQG 419
           |||||||:|||| |||||||||||||:|||||||| ||||  |:|| ||||| |:|
Db     347 RFGPDTQHLILNENCPSVHNLRSHKIQTQLSLIHPDIFPQLTSFYSKEEGSTLSVPTVRG 406

Qy     420 ECLLKYQLRPRREWQRDAIITCNPEEFIVEALQLPNFQQSVQEYRRSAQDGPAPAEKRSQ 479
           |||||| :||:||||||  : || :||| |||:||:||:||:|||:: |: |||||||||
Db     407 ECLLKYSVRPKREWQRDTTLDCNTDEFIAEALELPSFQESVEEYRKNVQENPAPAEKRSQ 466

Qy     480 YPEIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDCGEGTFGQLCRHYGDQVDRVLGTL 539
           |||||:|||||||||:|||||:||||:||| |:|||||||||||||| |:|||| :|
Db     467 YPEIVFLGTGSAIPMEIRNVSSTLVNLSPDKSVLLDCGEGTFGQLCRHYGQQIDRVLCSL 526

Qy     540 AAVFVSHLHADHHTGLPSILLQRERALASLGKPLHPLLVVAPNQLKAWLQQYHNQCQEVL 599
           ||||||||||||||| :||||||  |||||||  ||||||| ||:|||||| |||:||
Db     527 TAVFVSHLHADHHTGLLNILLQREHALASLGKPFQPLLVVAPTQLRAWLQQYHNHCQEIL 586

Qy     600 HHISMIPAKCLQEGAEISSPAVERLISSLLRTCDLEEFQTCLVRHCKHAFGCALVHTSGW 659
           ||:|||||||:|||:|:  :|||||  || |||||||||||||||||||||||:|||
Db     587 HHVSMIPAKCLQKGAEVSNTTLERLISLLLETCDLEEFQTCLVRHCKHAFGCALVHSSGW 646

Qy     660 KVVYSGDTMPCEALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMRMNAEF 719
           |||||||||||||||:|||||||||||||||| |||||||:|||||||:|||||||||||
Db     647 KVVYSGDTMPCEALVQMGKDATLLIHEATLEDXLEEEAVERTHSTTSQAINVGMRMNAEF 706

Qy     720 IMLNHFSQRYAKVPLFSPNFSEKVGVAFDHMKVCFGDFPTMPKLIPPLKALFAGDIEEME 779
           |||||||||| |:|||||:|:||||:|||||| |||||||:|||||||||||||||||||
Db     707 IMLNHFSQRYXKIPLFSPDFNEKVGIAFDHMKVXFGDFPTVPKLIPPLKALFAGDIEEMV 766

Qy     780 ERREKRELRQVRAALLSRELAGGLEDGEPQQKRAHTEE---PQAKK 822
```

```
Db    767 ERREKRELRLVRAALLTQQ-ADSPEDREPQQKRAHTDEPHSPQSKK 811


RESULT 8
US-09-988-687-222
; Sequence 222, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
;  APPLICANT: Tavtigian, Sean V.
;  APPLICANT:  Teng, David H.F.
;  APPLICANT:  Simard, Jacques
;  APPLICANT:  Rommens, Johanna M.
;  APPLICANT:  Myriad Genetics, Inc.
;  TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
;  TITLE OF INVENTION:  Gene and a Paralog and Orthologous Genes
;  FILE REFERENCE: 2318-258
;  CURRENT APPLICATION NUMBER: US/09/988,687
;  CURRENT FILING DATE:  2001-11-20
;  PRIOR APPLICATION NUMBER: 09/564,805
;  PRIOR FILING DATE: 2000-05-05
;  PRIOR APPLICATION NUMBER: US 60/107,468
;  PRIOR FILING DATE: 1998-11-06
;  PRIOR APPLICATION NUMBER: 09/434,382
;  PRIOR FILING DATE: 1999-11-05
;  NUMBER OF SEQ ID NOS: 240
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 222
;   LENGTH: 822
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-988-687-222

  Query Match             80.3%;  Score 3473.5;  DB 9;  Length 822;
  Best Local Similarity   80.5%;  Pred. No. 6.9e-294;
  Matches  665;  Conservative   66;  Mismatches   76;  Indels   19;  Gaps    6;

Qy      1 MWALCSLLRSAAGRTMSQGRTISQAPARRERPRKDPLRHLRTREKRGPSGCSGGPNTVYL 60
Db      1 MWALRSLLRPLGLRTMSQG------SARRPRPPKDPLRHLRTREKRGPG--PGGPNTVYL 52

Qy     61 QVVAAGSRDSGAALYVFSEFNRYLFNCGEGVQRLMQEHKLKVARLDNIFLTRMHWSNVGG 120
Db     53 QVVAAGGRDAGAALYVFSEYNRYLFNCGEGVQRLMQEHKTESRSLDNIFLTRMHWSNVGG 112

Qy    121 LSGMILTLKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPEYEDETMTV 180
Db    113 LCGMILTLKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPEYKDETMTV 172

Qy    181 YQIPIHSEQRRGKHQPWQSPERPLSRLSPERSSDSESNENEPHLPHGVSQRRGV-RDSSL 239
Db    173 YQVPIHSERRCGKQQPSQSPRTSPNRLSPKQSSDSGSAEN------GQCQQESMGQGPSL 226

Qy    240 VVAFICKLHLKRGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSITHEGREILAEELCTP 299
Db    227 VVAFVCKLHLRKGNFLVLKAKELGLPVGTAAIAPIIAAVKDGKSITYEGREIAAEELCTP 286

Qy    300 PDPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPASVLVDSRYQQWME 359
Db    287 PDPGLVFIVVECPDEGFILPICENDTFKRYQAEADAPVALVVHIAPESVLIDSRYQQWME 346

Qy    360 RFGPDTQHLVLNENCASVHNLRSHKIQTQLNLIHPDIFPLLTSFRCKKEGPTLSVPMVQG 419
Db    347 RFGPDTQHLILNENCPSVHNLRSHKIQTQLSLIHPDIFPQLTSFYSKEEGSTLSVPTVRG 406

Qy    420 ECLLKYQLRPRREWQRDAIITCNPEEFIVEALQLPNFQQSVQEYRRSAQDGPAPAEKRSQ 479
Db    407 ECLLKYSVRPKREWQRDTTLDCNTDEFIAEALELPSFQESVEEYRKNVQENPAPAEKRSQ 466

Qy    480 YPEIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDCGEGTFGQLCRHYGDQVDRVLGTL 539
Db    467 YPEIVFLGTGSAIPMEIRNVSSTLVNLSPDKSVLLDCGEGTFGQLCRHYGQQIDRVLCSL 526
```

```
Qy    540 AAVFVSHLHADHHTGLPSILLQRERALASLGKPLHPLLVVAPNQLKAWLQQYHNQCQEVL 599
Db    527 TAVFVSHLHADHHTGLLNILLQREHALASLGKPFQPLLVVAPTQLRAWLQQYHNHCQEIL 586

Qy    600 HHISMIPAKCLQEGAEISSPAVERLISSLLRTCDLEEFQTCLVRHCKHAFGCALVHTSGW 659
Db    587 HHVSMIPAKCLQKGAEVSNTTLERLISLLLETCDLEEFQTCLVRHCKHAFGCALVHSSGW 646

Qy    660 KVVYSGDTMPCEALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMRMNAEF 719
Db    647 KVVYSGDTMPCEALVQMGKDATLLIHEATLEDXLEEEAVERTHSTTSQAINVGMRMNAEF 706

Qy    720 IMLNHFSQRYAKVPLFSPNFSEKVGVAFDHMKVCFGDFPTMPKLIPPLKALFAGDIEEME 779
Db    707 IMLNHFSQRYXKIPLFSPDFNEKVGIAFDHMKVXFGDFPTVPKLIPPLKALFAGDIEEMV 766

Qy    780 ERREKRELRQVRAALLSRELAGGLEDGEPQQKRAHTEE---PQAKK 822
Db    767 ERREKRELRLVRAALLTQQ-ADSPEDREPQQKRAHTDEPHSPQSKK 811


RESULT 9
US-09-988-626-228
; Sequence 228, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
;  APPLICANT: Tavtigian, Sean V.
;  APPLICANT:  Teng, David H.F.
;  APPLICANT:  Simard, Jacques
;  APPLICANT:  Rommens, Johanna M.
;  APPLICANT:  Myriad Genetics, Inc.
;  TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
;  TITLE OF INVENTION:  Gene and a Paralog and Orthologous Genes
;  FILE REFERENCE: 2318-258
;  CURRENT APPLICATION NUMBER: US/09/988,626
;  CURRENT FILING DATE:  2001-11-20
;  PRIOR APPLICATION NUMBER: 09/564,805
;  PRIOR FILING DATE: 2000-05-05
;  PRIOR APPLICATION NUMBER: US 60/107,468
;  PRIOR FILING DATE: 1998-11-06
;  PRIOR APPLICATION NUMBER: 09/434,382
;  PRIOR FILING DATE: 1999-11-05
;  NUMBER OF SEQ ID NOS: 240
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 228
;   LENGTH: 837
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-09-988-626-228

  Query Match             20.2%;  Score 875.5;  DB 9;  Length 837;
  Best Local Similarity   29.4%;  Pred. No. 2.4e-67;
  Matches  250;  Conservative  128;  Mismatches  297;  Indels  175;  Gaps   28;

Qy     41 RTREKRGPSGCSGGPNTV-YLQVVAAG--SRDSGAALYVFSEFNRYLFNCGEGVQRLMQE 97
Db     39 RKSQKLNPT------NTIAYAQILGTGMDTQDTSSSVLLFFDKQRFIFNAGEGLQRFCTE 92

Qy     98 HKLKVARLDNIFLTRMHWSNVGGLSGMILTLK---ETGLPKCVLSGPPQLEKYLEAIKIF 154
Db     93 HKIKLSKIDHVFLSRVCSETAGGLPGLLLTLAGIGEEGLSVNVW-GPSDLNYLVDAMKSF 151

Qy    155 SGPLKGIEL-AVRPHSAPE----YEDETMTVYQI---PIHSEQRRGKHQPWQSPERPLSR 206
Db    152 IPRAAMVHTRSFGPSSTPDPIVLVNDEVVKISAIILKPCHSEE---------------- 194

Qy    207 LSPERSSDSESNENEPHLPHGVSQRRGVRDSSLVVAFICKLHLKRGNFLVLKAKEM-GLP 265
Db    195 -------DS----------------GNKSGDLSVVYVCELPEILGKFDLEKAKKVFGVK 230

Qy    266 VGTAAIAPIIAAVKDGKSITHEGREILA--EELCTPPDPGAAFVVVECPDESFIQPICEN 323
Db    231 PG-----PKYSRLQSGESVKSDERDITVHPSDVMGPSLPGPIVLLVDCPTESHAAELFSL 285
```

```
Qy    324 ATFQRYQGKADAP------VALVVHMAPASVLVDSRYQQWMERFGPDTQHLV-------- 369
          : : |    |        |  ::|::|:||     || ||::|   |||::
Db    286 KSLESYYSSPDEQTIGAKFVNCIIHLSPSSVTSSPTYQSWMKKFHL-TQHILAGHQRFLP 344

Qy    370 ----------LNENCASVHNLRSHKIQTQLNLIHPDIFPLLTSFRCKKEGPTLSVPMVQG 419
                   : :| |      | :|  :|| : |  ||   :  :    ::  |
Db    345 LLIIVSHQKTVRKNMAFPILKASSRIAARLNYLCPQFFPAPGFWPSQLTDNSIIDPTPSN 404

Qy    420 ECLLKYQLRP--RREWQRDAIITCNPEEFIVEAL--QLPNFQQSVQEYRR--SAQDGPAP 473
          |: |||    |   |  |       :|: |  ::|  :   :| ::  : |
Db    405 ----KFNLRPVAIRGIDRSCIPAPLTSSEVVDELLSEIPEIKDKSEEIKQFWNKQHNKTI 460

Qy    474 AEK-----------------RSQYPEIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDC 516
           ||                 |    ||: |||||: | | |||||  :::    ||||||
Db    461 IEKLWLSECNTVLPNCLEKIRRDDMEIVILGTGSSQPSKYRNVSAIFIDLFSRGSLLLDC 520

Qy    517 GEGTFGQLCRHYG-DQVDRVLGTLAAVFVSHLHADHHTGLPSILLQRERALASLGKPLHP 575
          ||||| ||| | || |  |  :  |  :::||:|||||||  ||  | : |   |    |
Db    521 GEGTLGQLKRRYGLDGADEAVRKLRCIWISHIHADHHTGLARILALRSKLLK--GVTHEP 578

Qy    576 LLVVAPNQLKAWLQQYHNQCQEVLHHISMIPAKC----------LQEGAEI-------SS 618
          ::|| |  || :|  |       |  : |    |           |: | |         |
Db    579 VIVVGPRPLKRFLDAYQR-----LEDLDMEFLDCRSTTATSWASLESGGEAEGSLFTQGS 633

Qy    619 PAVE-----------------RLISSLLRTCDLEEFQTCLVRHCKHAFGCALVHTS--- 657
          |                    : :  :|    | :  :  | ||  |:|  :
Db    634 PMQSVFKRSDISMDNSSVLLCLKNLKKVLSEIGLNDLISFPVVHCPQAYGVVIKAAERVN 693

Qy    658 -------GWKVVYSGDTMPCEALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAIS 710
                ||||:|||||: ||   |   :|||:|||||| || | |||: | ||||:||
Db    694 SVGEQILGWKMVYSGDSRPCPETVEASRDATILIHEATFEDALIEEALAKNHSTTKEAID 753

Qy    711 VGMRMNAEFIMLNHFSQRYAKVPLFSPNFSEKVGVAFDHMKVCFGDFPTMPKLIPPLKAL 770
          ||   |   |:| |||||| |:|:   :      :||| | :   |  :||::|  | |
Db    754 VGSAANVYRIVLTHFSQRYPKIPVIDESHMHNTCIAFDLMSINMADLHVLPKVLPYFKTL 813

Qy    771 FAGDIEEMEE 780
          |  :: | |:
Db    814 FRDEMVEDED 823
```

RESULT 10
US-09-988-687-228
; Sequence 228, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 228
; LENGTH: 837
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-988-687-228

Query Match 20.2%; Score 875.5; DB 9; Length 837;
Best Local Similarity 29.4%; Pred. No. 2.4e-67;

Matches 250; Conservative 128; Mismatches 297; Indels 175; Gaps 28;

```
Qy     41 RTREKRGPSGCSGGPNTV-YLQVVAAG--SRDSGAALYVFSEFNRYLFNCGEGVQRLMQE 97
          |  :|  |:      ||: | |::  |  ::|: ::: :| :  |::|| |||:||   |
Db     39 RKSQKLNPT------NTIAYAQILGTGMDTQDTSSSVLLFFDKQRFIFNAGEGLQRFCTE 92

Qy     98 HKLKVARLDNIFLTRMHWSNVGGLSGMILTLK---ETGLPKCVLSGPPQLEKYLEAIKIF 154
          ||:|::::|::||:|:|:     ||| |::|||    | ||   |  ||  |   ::|:| |
Db     93 HKIKLSKIDHVFLSRVCSETAGGLPGLLLTLAGIGEEGLSVNVW-GPSDLNYLVDAMKSF 151

Qy    155 SGPLKGIEL-AVRPHSAPE----YEDETMTVYQI---PIHSEQRRGKHQPWQSPERPLSR 206
                :   : | | |:      || : :  |   | |||:
Db    152 IPRAAMVHTRSFGPSSTPDPIVLVNDEVVKISAIILKPCHSEE----------------- 194

Qy    207 LSPERSSDSESNENEPHLPHGVSQRRGVRDSSLVVAFICKLHLKRGNFLVLKAKEM-GLP 265
                ||                  | :   | | ::|:|    | | : |||:: |:
Db    195 -------DS----------------GNKSGDLSVVYVCELPEILGKFDLEKAKKVFGVK 230

Qy    266 VGTAAIAPIIAAVKDGKSITHEGREILA--EELCTPPDPGAAFVVVECPDESFIQPICEN 323
          |      |  : :: |:|:  : |:|      ::  |  ||   ::|:|| ||    :
Db    231 PG-----PKYSRLQSGESVKSDERDITVHPSDVMGPSLPGPIVLLVDCPTESHAAELFSL 285

Qy    324 ATFQRYQGKADAP------VALVVHMAPASVLVDSRYQQWMERFGPDTQHLV-------- 369
          : : |    |        |  ::|::|:|     || ||::|   |||::
Db    286 KSLESYYSSPDEQTIGAKFVNCIIHLSPSSVTSSPTYQSWMKKFHL-TQHILAGHQRFLP 344

Qy    370 ----------LNENCASVHNLRSHKIQTQLNLIHPDIFPLLTSFRCKKEGPTLSVPMVQG 419
                   : :| |      | :|  :|| : |  ||   :  :    ::  |
Db    345 LLIIVSHQKTVRKNMAFPILKASSRIAARLNYLCPQFFPAPGFWPSQLTDNSIIDPTPSN 404

Qy    420 ECLLKYQLRP--RREWQRDAIITCNPEEFIVEAL--QLPNFQQSVQEYRR--SAQDGPAP 473
          |: |||    |   |  |       :|: |  ::|  :   :| ::  : |
Db    405 ----KFNLRPVAIRGIDRSCIPAPLTSSEVVDELLSEIPEIKDKSEEIKQFWNKQHNKTI 460

Qy    474 AEK-----------------RSQYPEIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDC 516
           ||                 |    ||: |||||: | | |||||  :::    ||||||
Db    461 IEKLWLSECNTVLPNCLEKIRRDDMEIVILGTGSSQPSKYRNVSAIFIDLFSRGSLLLDC 520

Qy    517 GEGTFGQLCRHYG-DQVDRVLGTLAAVFVSHLHADHHTGLPSILLQRERALASLGKPLHP 575
          ||||| ||| | || |  |  :  |  :::||:|||||||  ||  | : |   |    |
Db    521 GEGTLGQLKRRYGLDGADEAVRKLRCIWISHIHADHHTGLARILALRSKLLK--GVTHEP 578

Qy    576 LLVVAPNQLKAWLQQYHNQCQEVLHHISMIPAKC----------LQEGAEI-------SS 618
          ::|| |  || :|  |       |  : |    |           |: | |         |
Db    579 VIVVGPRPLKRFLDAYQR-----LEDLDMEFLDCRSTTATSWASLESGGEAEGSLFTQGS 633

Qy    619 PAVE-----------------RLISSLLRTCDLEEFQTCLVRHCKHAFGCALVHTS--- 657
          |                    : :  :|    | :  :  | ||  |:|  :
Db    634 PMQSVFKRSDISMDNSSVLLCLKNLKKVLSEIGLNDLISFPVVHCPQAYGVVIKAAERVN 693

Qy    658 -------GWKVVYSGDTMPCEALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAIS 710
                ||||:|||||: ||   |   :|||:|||||| || | |||: | ||||:||
Db    694 SVGEQILGWKMVYSGDSRPCPETVEASRDATILIHEATFEDALIEEALAKNHSTTKEAID 753

Qy    711 VGMRMNAEFIMLNHFSQRYAKVPLFSPNFSEKVGVAFDHMKVCFGDFPTMPKLIPPLKAL 770
          ||   |   |:| |||||| |:|:   :      :||| | :   |  :||::|  | |
Db    754 VGSAANVYRIVLTHFSQRYPKIPVIDESHMHNTCIAFDLMSINMADLHVLPKVLPYFKTL 813

Qy    771 FAGDIEEMEE 780
          |  :: | |:
Db    814 FRDEMVEDED 823
```

RESULT 11
US-09-988-626-227
; Sequence 227, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.

; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 227
; LENGTH: 844
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-988-626-227

Query Match 17.6%; Score 760; DB 9; Length 844;
Best Local Similarity 26.6%; Pred. No. 2.8e-57;
Matches 226; Conservative 175; Mismatches 324; Indels 126; Gaps 29;

```
Qy       9 RSAAGRTMSQGRTISQAPARRERPRKDPLRH----LRTREKRGPSGCSGG----PNTVYL 60
Db      36 RIARNRRILQKHSSSHLKAREVNASISNLRQSMAAVQKKQKAAHEPPANSIVNIPSQVSI 95

Qy      61 QVVAAGSRDSGAALYVFSEFNRYLFNCGEGVQRLMQEHKLKVARLDNIFLTRMHWSNVGG 120
Db      96 EVLGNGTGLLRACFILRTPLKTYMFNCPENACRFLWQLRIRSSSVVDLFITSANWDNIAG 155

Qy     121 LSGMILTLKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKG-----IELAVRPHSAPEYED 175
Db     156 ISSILLS-KESNALSTRLHGAMNIKHFLECIRPFQDSDYGSCKYPSQVEERPYTMENYED 214

Qy     176 ETMTVYQIPIHSEQRRGKHQPWQSPERPLSRLSPERSSDSESNENEPHLPHGVSQRRGVR 235
Db     215 AGLKVTYIPL------------SP--PLN-----IGSNNEKSKN------------VK 241

Qy     236 DSSLVVAFICKLHLKRGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSIT-HEGREILAE 294
Db     242 VNNVDIAFLIEMKEAARRIDTMKLMELKVPKG-----PLIGKLKSGEAVTLPDGRTIQPD 296

Qy     295 ELCTP---PDPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPASVLVD 351
Db     297 QVFSSDKVEGDKPLLLVTECTTEDHVKALIDSSSLQPFL-NGEKQLDYMVHISDDAVINT 355

Qy     352 SRYQQWMERF-GPDTQHLVLNENCASVHNLRS-HKIQTQLNLIHPDIFPLLTSFRCKKEG 409
Db     356 PTYRHLMEKLNNPSITHLLINGGNPVIPAVESVYKHTRLLRSIAPSLFPALHPI------ 409

Qy     410 PTLSVPMVQGECLLKYQ-----LRP-RREWQRDAIITCNPEEFIVEALQL-----PNFQQ 458
Db     410 -DWSGIITQNEELSQRQDQFIRVAPMQRYWMRRG-ASFNEEPIVNNLLAAEPELSDKAKE 467

Qy     459 SVQEYRRSAQDGPAPAEKRSQYPEIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDCGE 518
Db     468 LIKEYQKLEKENKMDCE----FPKLTFFGTSSAVPSKYRNVTGYLVEASENSAILIDVGE 523

Qy     519 GTFGQLCRHYG-DQVDRVLGTLAAVFVSHLHADHHTGLPSILLQRERALASLGKPLHPLL 577
Db     524 GTYGQMRAVFGEDGCKQLLVNLNCVLITHAHQDHMNGLYTIIARRKEAFESLGAPYRPLV 583

Qy     578 VVAPNQLKAWLQQYHNQCQEVLHHISMI------------PAKCLQEGAEISSP----- 619
Db     584 LVCNRNVLKPMKTY-SICFENIEHLLEIVDISRYPLTPPGSPGGPPGKRPRLPSPHLPPS 642

Qy     620 --AVERLISSL-LRTCDLEEFQTCLVRHCKHAFGCALVHTSGWKVVYSGDTMPCEALVRM 676
Db     643 RDVLQDMSSSFDKKAWKLDELKAVQVHHTRMANG-FVMRVAGKRIVFSGDTKPCDLLVEE 701

Qy     677 GKDATLLIHEATLEDGLE----------------EEAVEKTHSTTSQAISVGMRMNAEF 719
Db     702 GKDADVLVHESTFEDGHEVDMTPKPPKKLAKISSLADAMRKRHSTMGQAVDVGKRMNAKH 761

Qy     720 IMLNHFSQRYAKVPLFSPNF--SEKVGVAFDHMKVCFGDFPTMPKLIPPLKALFAGDIEE 777
Db     762 IILTHFSARYPKVPVL-PEYLDKENIGVAMDMLRVRFDHLPLVSKLLPIFREVFVAELFE 820

Qy     778 MEERREKRELR 788
Db     821 LTIKKEQRVLK 831
```

RESULT 12
US-09-988-687-227
; Sequence 227, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 227
; LENGTH: 844
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-988-687-227

Query Match 17.6%; Score 760; DB 9; Length 844;
Best Local Similarity 26.6%; Pred. No. 2.8e-57;
Matches 226; Conservative 175; Mismatches 324; Indels 126; Gaps 29;

```
Qy       9 RSAAGRTMSQGRTISQAPARRERPRKDPLRH----LRTREKRGPSGCSGG----PNTVYL 60
Db      36 RIARNRRILQKHSSSHLKAREVNASISNLRQSMAAVQKKQKAAHEPPANSIVNIPSQVSI 95

Qy      61 QVVAAGSRDSGAALYVFSEFNRYLFNCGEGVQRLMQEHKLKVARLDNIFLTRMHWSNVGG 120
Db      96 EVLGNGTGLLRACFILRTPLKTYMFNCPENACRFLWQLRIRSSSVVDLFITSANWDNIAG 155

Qy     121 LSGMILTLKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKG-----IELAVRPHSAPEYED 175
Db     156 ISSILLS-KESNALSTRLHGAMNIKHFLECIRPFQDSDYGSCKYPSQVEERPYTMENYED 214

Qy     176 ETMTVYQIPIHSEQRRGKHQPWQSPERPLSRLSPERSSDSESNENEPHLPHGVSQRRGVR 235
Db     215 AGLKVTYIPL------------SP--PLN-----IGSNNEKSKN------------VK 241

Qy     236 DSSLVVAFICKLHLKRGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSIT-HEGREILAE 294
Db     242 VNNVDIAFLIEMKEAARRIDTMKLMELKVPKG-----PLIGKLKSGEAVTLPDGRTIQPD 296

Qy     295 ELCTP---PDPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPASVLVD 351
Db     297 QVFSSDKVEGDKPLLLVTECTTEDHVKALIDSSSLQPFL-NGEKQLDYMVHISDDAVINT 355

Qy     352 SRYQQWMERF-GPDTQHLVLNENCASVHNLRS-HKIQTQLNLIHPDIFPLLTSFRCKKEG 409
Db     356 PTYRHLMEKLNNPSITHLLINGGNPVIPAVESVYKHTRLLRSIAPSLFPALHPI------ 409

Qy     410 PTLSVPMVQGECLLKYQ-----LRP-RREWQRDAIITCNPEEFIVEALQL-----PNFQQ 458
```

```
Db    410 -DWSGIITQNEELSQRQDQFIRVAPMQRYWMRRG-ASFNEEPIVNNLLAAEPELSDKAKE 467

Qy    459 SVQEYRRSAQDGPAPAEKRSQYPEIIFLGTGSAIPMKIRNVSATLVNISPDTSLLLDCGE 518

Db    468 LIKEYQKLEKENKMDCE----FPKLTFFGTSSAVPSKYRNVTGYLVEASENSAILIDVGE 523

Qy    519 GTFGQLCRHYG-DQVDRVLGTLAAVFVSHLHADHHTGLPSILLQRERALASLGKPLHPLL 577

Db    524 GTYGQMRAVFGEDGCKQLLVNLNCVLITHAHQDHMNGLYTIIARRKEAFESLGAPYRPLV 583

Qy    578 VVAPNQLKAWLQQYHNQCQEVLHHISMI------------PAKCLQEGAEISSP----- 619

Db    584 LVCNRNVLKPMKTY-SICFENIEHLLEIVDISRYPLTPPGSPGGPPGKRPRLPSPHLPPS 642

Qy    620 --AVERLISSL-LRTCDLEEFQTCLVRHCKHAFGCALVHTSGWKVVYSGDTMPCEALVRM 676

Db    643 RDVLQDMSSSFDKKAWKLDELKAVQVHHTRMANG-FVMRVAGKRIVFSGDTKPCDLLVEE 701

Qy    677 GKDATLLIHEATLEDGLE---------------EEAVEKTHSTTSQAISVGMRMNAEF 719

Db    702 GKDADVLVHESTFEDGHEVDMTPKPPKKLAKISSLADAMRKRHSTMGQAVDVGKRMNAKH 761

Qy    720 IMLNHFSQRYAKVPLFSPNF--SEKVGVAFDHMKVCFGDFPTMPKLIPPLKALFAGDIEE 777

Db    762 IILTHFSARYPKVPVL-PEYLDKENIGVAMDMLRVRFDHLPLVSKLLPIFREVFVAELFE 820

Qy    778 MEERREKRELR 788

Db    821 LTIKKEQRVLK 831


RESULT 13
US-09-988-626-229
; Sequence 229, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
;  APPLICANT: Tavtigian, Sean V.
;  APPLICANT:  Teng, David H.F.
;  APPLICANT:  Simard, Jacques
;  APPLICANT:  Rommens, Johanna M.
;  APPLICANT:  Myriad Genetics, Inc.
;  TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
;  TITLE OF INVENTION:  Gene and a Paralog and Orthologous Genes
;  FILE REFERENCE: 2318-258
;  CURRENT APPLICATION NUMBER: US/09/988,626
;  CURRENT FILING DATE:  2001-11-20
;  PRIOR APPLICATION NUMBER: 09/564,805
;  PRIOR FILING DATE: 2000-05-05
;  PRIOR APPLICATION NUMBER: US 60/107,468
;  PRIOR FILING DATE: 1998-11-06
;  PRIOR APPLICATION NUMBER: 09/434,382
;  PRIOR FILING DATE: 1999-11-05
;  NUMBER OF SEQ ID NOS: 240
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 229
;   LENGTH: 838
;   TYPE: PRT
;   ORGANISM: Saccharomyces cerevisiae
US-09-988-626-229

  Query Match             13.9%;  Score 599.5;  DB 9;  Length 838;
  Best Local Similarity   25.7%;  Pred. No. 2.8e-43;
  Matches  221;  Conservative  138;  Mismatches  290;  Indels  211;  Gaps   36;

Qy     82 RYLF-NCGEGVQRLMQEHKLKVARLDNIFLT-RMHWSNVGGLSGMILTLKETGLPKCVLS 139

Db     28 KYFFGKIGEGSQRSLTENKIRISKLKDIFLTGELNWSDIGGLPGMILTIADQGKSNLVLH 87

Qy    140 GPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPE---YEDETMTVYQIPI---HSEQRRGK 193

Db     88 YGNDILNYIVSTWRYFVFRFGIDL--NDHIMKDKEVYKDKIIAVKSFNVLKNGGEDRLGV 145

Qy    194 HQPWQS------------PERPLSRLSPERSSDSESNENEPHLPHGVSQRRGVRDSSLVV 241

Db    146 FDSFQKGVLRSIVAKMFPKHAPTDRYDP--SSDPHLNVELPDL-----------DAKVEV 192

Qy    242 AFICKLHLK--RGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSIT-HEGREILAEELCT 298

Db    193 STNYEISFSPVRGKFKVEEAIKLGVPKG-----PLFAKLTKGQTITLDNGIVVTPEQVLE 247

Qy    299 PPDPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPASVLVDSRYQQWM 358

Db    248 NERHFAKVLILDIPDDLYL-----NAFVEKFKDYDCAELGMVYYFLGDEVTINDNLFAFI 302

Qy    359 ERFGPDTQHLVLNENCASVHNLRSHKIQTQLNLIHPDIFPL----LTSFRCK-------- 406

Db    303 DIFE--------KNNYGKVNHMISH------NKISPNTISFFGSALTTLKLKALQVNNYN 348

Qy    407 --KEGPTLS--------VPMVQGECLLKYQLRPRRE---------WQRDAIITCNP---- 443

Db    349 LPKTDRVFSKDFYDRFDTPLSRGTSMCKSQEEPLNTIIEKDNIHIFSQNKTVTFEPFRMN 408

Qy    444 ----------------------EEFIVEALQLP--------NFQQSVQEYRRSAQDGPAPA 474

Db    409 EEPMKCNINGEVADFSWQEIFEEH-VKPLEFPLADVDTVINNQLHVDNFNNSAE------ 461

Qy    475 EKRSQYPEIIFLGTGSAIPMKIRNVSATLVNI--------SPDTSLLLDCGEGTFGQLCR 526

Db    462 --KKKHVEIITLGTGSALPSKYRNVVSTLVKVPFTDADGNTINRNIMLDAGENTLGTIHR 519

Qy    527 HYGD-QVDRVLGTLAAVFVSHLHADHHTGLPSILLQRERALASLGKPLHPLLVVAPNQLK 585

Db    520 MFSQLAVKSIFQDLKMIYLSHLHADHHLGIISVL--NEWYKYNKDDETSYIYVVTP---- 573

Qy    586 AWLQQYHN--------QCQEVLHHISMIPA----------------------KCLQEGA- 614

Db    574 -W--QYHKFVNEWLVLENKEILKRIKYISCEHFINDSFVRMQTQSVPLAEFNEILKENSN 630

Qy    615 -------EISSPAVER---LISSLLRTCDLEEFQTCLVRHCKHAFGCALV--------HT 656

Db    631 QESNRKLELDRDSSYRDVDLIRQMYEDLSIEYFQTCRAIHCDWAYSNSITFRMDENNEHN 690

Qy    657 SGWKVVYSGDTMPC--EALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMR 714

Db    691 T-FKVSYSGDTRPNIEKFSLEIGYNSDLLIHEATLENQLLEDAVKKKHCTINEAIGVSNK 749

Qy    715 MNAEFIMLNHFSQRYAKVPLFSPN---FSEKVGVAFDHMKVCFGDFPTMPKLIPPLKALF 771

Db    750 MNARKLILTHFSQRYPKLPQLDNNIDVMAREFCFAFDSMIVDYEKIGEQQRIFPLLNKAF 809

Qy    772 AGDIEEMEERREKRELRQVR 791

Db    810 ---VEEKEEEEDVDDVESVQ 826


RESULT 14
US-09-988-687-229
; Sequence 229, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
;  APPLICANT: Tavtigian, Sean V.
;  APPLICANT:  Teng, David H.F.
;  APPLICANT:  Simard, Jacques
;  APPLICANT:  Rommens, Johanna M.
;  APPLICANT:  Myriad Genetics, Inc.
;  TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
;  TITLE OF INVENTION:  Gene and a Paralog and Orthologous Genes
;  FILE REFERENCE: 2318-258
;  CURRENT APPLICATION NUMBER: US/09/988,687
;  CURRENT FILING DATE:  2001-11-20
;  PRIOR APPLICATION NUMBER: 09/564,805
;  PRIOR FILING DATE: 2000-05-05
;  PRIOR APPLICATION NUMBER: US 60/107,468
;  PRIOR FILING DATE: 1998-11-06
;  PRIOR APPLICATION NUMBER: 09/434,382
```

```
;  PRIOR FILING DATE: 1999-11-05
;  NUMBER OF SEQ ID NOS: 240
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 229
;   LENGTH: 838
;   TYPE: PRT
;   ORGANISM: Saccharomyces cerevisiae
US-09-988-687-229

  Query Match             13.9%;  Score 599.5;  DB 9;  Length 838;
  Best Local Similarity   25.7%;  Pred. No. 2.8e-43;
  Matches  221;  Conservative  138;  Mismatches  290;  Indels  211;  Gaps   36;

Qy      82 RYLF-NCGEGVQRLMQEHKLKVARLDNIFLT-RMHWSNVGGLSGMILTLKETGLPKCVLS 139
           :| |   ||| || : |:|:::::| :|||| ::||::||| |||||: : |    ||
Db      28 KYFFGKIGEGSQRSLTENKIRISKLKDIFLTGELNWSDIGGLPGMILTIADQGKSNLVLH 87

Qy     140 GPPQLEKYLEAIKIFSGPLKGIELAVRPHSAPE---YEDETMTVYQIPI---HSEQRRGK 193
             :  |: :    :     ||:|    |   :   |:|: : |    :     | | |
Db      88 YGNDILNYIVSTWRYFVFRFGIDL--NDHIMKDKEVYKDKIIAVKSFNVLKNGGEDRLGV 145

Qy     194 HQPWQS------------PERPLSRLSPERSSDSESNENEPHLPHGVSQRRGVRDSSLVV 241
             :|                |  |  |  |||   |   | |           |: : |
Db     146 FDSFQKGVLRSIVAKMFPKHAPTDRYDP--SSDPHLNVELPDL-----------DAKVEV 192

Qy     242 AFICKLHLK--RGNFLVLKAKEMGLPVGTAAIAPIIAAVKDGKSIT-HEGREILAEELCT 298
           :   ::     || | | :| ::|:| |     |: | :  |::||   |  :  |::
Db     193 STNYEISFSPVRGKFKVEEAIKLGVPKG-----PLFAKLTKGQTITLDNGIVVTPEQVLE 247

Qy     299 PPDPGAAFVVVECPDESFIQPICENATFQRYQGKADAPVALVVHMAPASVLVDSRYQQWM 358
             |  :::: ||: ::     ||  ::::    | : :| :     | ::     ::
Db     248 NERHFAKVLILDIPDDLYL-----NAFVEKFKDYDCAELGMVYYFLGDEVTINDNLFAFI 302

Qy     359 ERFGPDTQHLVLNENCASVHNLRSHKIQTQLNLIHPDIFPL----LTSFRCK-------- 406
           : |           |   |::: ||      | | |:        ||: : |
Db     303 DIFE--------KNNYGKVNHMISH------NKISPNTISFFGSALTTLKLKALQVNNYN 348

Qy     407 --KEGPTLS--------VPMVQGECLLKYQLRPRRE---------WQRDAIITCNP---- 443
             |     |          |: :|  : | |  |           : ::  :|  |
Db     349 LPKTDRVFSKDFYDRFDTPLSRGTSMCKSQEEPLNTIIEKDNIHIFSQNKTVTFEPFRMN 408

Qy     444 ---------------------EEFIVEALQLP--------NFQQSVQEYRRSAQDGPAPA 474
                               ||  |: |: |         | |  |  :  ||:
Db     409 EEPMKCNINGEVADFSWQEIFEEH-VKPLEFPLADVDTVINNQLHVDNFNNSAE------ 461

Qy     475 EKRSQYPEIIFLGTGSAIPMKIRNVSATLVNI--------SPDTSLLLDCGEGTFGQLCR 526
             : :: ||| |||||:| | ||| :||| :         : : :::|| || | | : |
Db     462 --KKKHVEIITLGTGSALPSKYRNVVSTLVKVPFTDADGNTINRNIMLDAGENTLGTIHR 519

Qy     527 HYGD-QVDRVLGTLAAVFVSHLHADHHTGLPSILLQRERALASLGKPLHPLLVVAPNQLK 585
            :    |  :   |  :::|||||||| |: |:|   |    :       : || |
Db     520 MFSQLAVKSIFQDLKMIYLSHLHADHHLGIISVL--NEWYKYNKDDETSYIYVVTP---- 573

Qy     586 AWLQQYHN--------QCQEVLHHISMIPA---------------------KCLQEGA- 614
             |  |||          : :|:|  |  |                     : |:| :
Db     574 -W--QYHKFVNEWLVLENKEILKRIKYISCEHFINDSFVRMQTQSVPLAEFNEILKENSN 630

Qy     615 -------EISSPAVER---LISSLLRTCDLEEFQTCLVRHCKHAFGCALV--------HT 656
                  |:   :  |   ||  :     :| ||||   ||  |:  ::          |
Db     631 QESNRKLELDRDSSYRDVDLIRQMYEDLSIEYFQTCRAIHCDWAYSNSITFRMDENNEHN 690

Qy     657 SGWKVVYSGDTMPC--EALVRMGKDATLLIHEATLEDGLEEEAVEKTHSTTSQAISVGMR 714
           : :|| ||||| |   :  : :| :: |||||||||: | |:||:| | | ::|| |  :
Db     691 T-FKVSYSGDTRPNIEKFSLEIGYNSDLLIHEATLENQLLEDAVKKKHCTINEAIGVSNK 749

Qy     715 MNAEFIMLNHFSQRYAKVPLFSPN---FSEKVGVAFDHMKVCFGDFPTMPKLIPPLKALF 771
           |||  ::| |||||| |:|    |    : :   ||| | | :       :: | |   |
Db     750 MNARKLILTHFSQRYPKLPQLDNNIDVMAREFCFAFDSMIVDYEKIGEQQRIFPLLNKAF 809

Qy     772 AGDIEEMEERREKRELRQVR 791
             :|| ||  :  ::  |:
Db     810 ---VEEKEEEEDVDDVESVQ 826
```

```
RESULT 15
US-09-988-626-211
; Sequence 211, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
;  APPLICANT: Tavtigian, Sean V.
;  APPLICANT:  Teng, David H.F.
;  APPLICANT:  Simard, Jacques
;  APPLICANT:  Rommens, Johanna M.
;  APPLICANT:  Myriad Genetics, Inc.
;  TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
;  TITLE OF INVENTION:  Gene and a Paralog and Orthologous Genes
;  FILE REFERENCE: 2318-258
;  CURRENT APPLICATION NUMBER: US/09/988,626
;  CURRENT FILING DATE:  2001-11-20
;  PRIOR APPLICATION NUMBER: 09/564,805
;  PRIOR FILING DATE: 2000-05-05
;  PRIOR APPLICATION NUMBER: US 60/107,468
;  PRIOR FILING DATE: 1998-11-06
;  PRIOR APPLICATION NUMBER: 09/434,382
;  PRIOR FILING DATE: 1999-11-05
;  NUMBER OF SEQ ID NOS: 240
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 211
;   LENGTH: 81
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-988-626-211

  Query Match             9.7%;  Score 420;  DB 9;  Length 81;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-29;
  Matches   81;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MWALCSLLRSAAGRTMSQGRTISQAPARRERPRKDPLRHLRTREKRGPSGCSGGPNTVYL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MWALCSLLRSAAGRTMSQGRTISQAPARRERPRKDPLRHLRTREKRGPSGCSGGPNTVYL 60

Qy      61 QVVAAGSRDSGAALYVFSEFN 81
           |||||||||||||||||||||
Db      61 QVVAAGSRDSGAALYVFSEFN 81


Search completed: May 14, 2003, 10:11:04
Job time : 30 secs
```